1. INTRODUCTION

One of the most frequently advocated incentives for the study of comets is that the cometary nuclei carry clues about the origin of the solar system. However, cometary observations almost never deal with the nucleus itself, but with its surrounding coma. An essential problem is therefore the derivation of nuclear properties from coma observations. In this chapter, we do not review the observations, nor the derived nuclear properties. Instead, we focus on the methods by which the latter are derived from the former: What are they? What are their formal justifications? Are the conclusions derived by these methods reliable? If not, which alternative methods should be used?

As discussed in detail in the literature (e.g., Weissman, 1999): (1) comets may have been formed from planetesimals over a large range of heliocentric distances, including the asteroid belt region, and therefore may differ from one another in composition and structure, in particular in volatile to refractory relative abundance; and (2) dynamical exchanges of material during the formation of the planetesimals probably mixed materials formed at different distances from the Sun, so that any cometary nucleus may be a mixture of materials formed at different temperatures, and therefore may be a complicated object. The shape of the nucleus should be expected to be as complicated as asteroidal shapes, hence its rotation should not be assumed to be simple (such as a symmetrical top rotator). Furthermore, while the nucleus interior is generally considered as undifferentiated, such cannot be the case of the external layers — those from which the coma is formed: They are submitted to thermal cycling (e.g., Espinasse et al., 1991), cosmic-ray irradiation (e.g., Strazzulla, 1997), and meteoroid impacts (Fernández, 1990). This intrinsic complexity suggests that the interpretation of coma observations in terms of nuclei properties cannot be trivial.

Ground-based coma observations typically provide images and spectra of the gas and dust coma at a spatial resolution not better than several hundreds of kilometers, and at a time resolution sometimes of hours but more typically of days. The overall appearance of the coma is usually simple, but image enhancement techniques reveal well-defined spatial structures both in the dust and in the gas coma. These structures often repeat themselves in time, quasiperiodically, and quasiperiodicity is also often observed in the global properties of the coma (superimposed upon the secular evolution due to the orbital motion). It has become standard practice to summarize the coma observations by phenomenological models. Here, as we shall demonstrate at length in the following sections of this chapter, (1) one assumes that the coma formed by the real, complicated nuclei is well understood after having discussed only the coma formed by the simplest conceivable nuclei (e.g., spherical in shape, emitting spherical dust grains, rotating uniformly, etc.); (2) even this simplest possible coma is not self-consistently modeled using the relevant well-established applied physics methods, but simply “understood” from subjective postulates (e.g., there exists rotationally-invariant gas and dust ejection patterns, “gas and dust heliocentric dependence velocity laws”, etc.) (see Crifo and Rodionov, 1999). This double arbitrariness in analyzing the data is perhaps tolerable as long as it is only used as a convenient means of summarizing observations, but, otherwise, it raises the question of whether the “derived” properties capture anything of the real properties of the real nuclei. This question can only be answered by taking into consideration realistically complicated nuclei, and thoroughly computing the structure of the coma they form — i.e., by physical modeling.

The 1986 flyby missions to Comet P/Halley have opened a new era. With the advent of observations that resolve the nucleus and sample the coma on a kilometer scale, a direct physical description becomes conceivable, i.e., one in which
the complexity of the object is taken into consideration in detail. But, in the same manner as radically new observational techniques are requested for the flyby observations, the objectives themselves of the observations and of their interpretation change dramatically. For instance, the technical success of future orbiting and landing mission itself requires that the immediate environment of the nucleus be forecasted with high reliability and with a quick reaction time from the data acquired from the probes, regardless of whether or not this gives “clues on the origin of comets.” This is a formidable challenge compared to that of interpreting grounded cometary observations.

Even when deep-space mission data are at hand, it is not proven that such data allow a thorough understanding of the nucleus: To investigate it involves a so-called “inverse problem” with an enormous number of unknowns. The only known approach to addressing such a problem is statistical simulation; the statistical properties of a large class of realistic (i.e., as complicated as possible) nuclei must be derived and compared to the observations.

Both short-term forecasting of a nucleus activity and statistical estimates of the significance of nucleus and coma observations require advanced physical modeling. We review here the few existing efforts that have been done in this direction. As deep-space missions to comets are scarce, the institutional support of these efforts is quite limited. In such a context, unfortunate habits persist across decades: One such habit is to take the phenomenological models for a true representation of the real objects, as if this did not raise any question; another one is to apply the “supposed approach” of (1) objects that plausibly represent real comets and (2) the simple fictitious objects usually considered, one can hope to separate the domain of robustly derived conclusions from the domain of unwarranted speculations.

The present chapter is organized as follows. In section 2, we illustrate by a few representative examples the underlying assumptions on which heuristic inferences have been made in the cometary literature. Sections 3–5 describe the existing physical models, placing emphasis on their physical significance, not on the mathematical methods. In section 6 we review the physical model results relative to the near-nucleus coma; special attention is given to whether they support the assumptions listed in section 2, or otherwise. Section 7 reviews the cases where model and observations are compared. Section 8 then briefly addresses the outer coma, again with an emphasis on the validity (or otherwise) of the phenomenological assumptions. Section 9 attempts to draw a general lesson from the previous sections.

2. OBSERVED COMA STRUCTURES AND THEIR HEURISTIC INTERPRETATIONS

By “structures,” we mean spatial details in the coma, or temporal patterns that can be recognized during the comet motion: Both are related, owing to the nucleus rotation. The status of interpretation of these structures just before the first elaborate physical model results appeared was reviewed in Kömle (1990). Here, we reexamine the former in the light of the latter. As many chapters of this book are dedicated to coma structures, we will present here only a few representative examples. Many are relative to Comet P/Halley because it has been the only comet whose full nucleus shape was derived from flyby images. (The January 2004 flyby observations of Comet P/Wild 2 will probably lead to the derivation of the full shape of a second nucleus.) A few other, most instructive examples are relative to Comets Swift-Tuttle, Hyakutake, and Hale-Bopp. When physical modeling is advocated in support of the interpretation of the structures, we review it carefully. Often, it is possible to point out uncertainties, misinterpretations, and sometimes severe inconsistencies in the heuristic interpretations or their supporting modeling, even before advocating physical model results: A critical internal consistency check of the approach is sufficient. Advanced physical models are unavoidable, however, to overcome the difficulties, as will be demonstrated.

It is also unavoidable to comment on the terminology often found in the literature: Words such as “jets,” “shells,” and “activity” are, as we shall demonstrate, often used either without their standard meaning as defined in physics, or sometimes even in conflict with it. In our opinion, this is extremely prejudicial.

2.1. Gas Light Curves

If one observes a comet with the same spectrophotometer at intervals of days or fractions of days, the molecular emissions are seen to exhibit secular variations related to changes in heliocentric distance $r_{\odot}$ plus, in general, short-term variations. A good example can be found in Millis and Schleicher (1986). Some of these fluctuations may be random (“bursts”), but more attention is usually given to those that suggest approximate periodicities, because of their unquestionable relation to the nucleus rotation.

With the exception of the peculiar case of Comet Hale-Bopp (see below), such short-term variations are observed at small heliocentric distances, where water production is dominant. Water molecules emissions from the coma are rarely observable, so that their production rate $Q(t)$ is usually derived from the number $N$ of daughter OH molecules present inside the observed coma volume (usually a cylinder). To relate $Q(t)$ to $N(t)$, observers use a trivial analytical
formula, the “Haser model” (the expression “Haser formula” would be more suitable). There are many recognized weaknesses in this formula. Let us only mention a fundamental one: The coma must be spherically symmetric and quasisteady. But the very interpretation of quasiperiodicities in gas lightcurves, in terms of nucleus rotation, implies that this is not the case. Due allowance for asphericity of the coma is therefore needed. Furthermore, since inside a very narrow field of view, molecules are present to very large distances (on the order of 10<sup>5</sup> km), the observational volume filling time is on the order of a day, comparable to typical rotation periods; there is “hysteresis,” i.e., a time-dependent model of the coma content must be used. This, in turn, requires knowing the nucleus rotational state.

Some authors are careful to point out that their use of the Haser formula is only done for convenience, the accuracy of the algorithm being unknown. Our impression is, however, that many readers take the Haser-derived Q(t) as a measure of the algorithm being unknown. Our interpretation is, the Haser formula is only done for convenience, the accuracy of the algorithm being unknown.

2. Comet P/Halley Gas “Jets” and “Shells”

Both ground-based and flyby spacecraft data revealed that the gas coma of P/Halley is structured. Clairemidi et al. (1990) identified OH and NH brightness maxima, which they called “jets,” in the images obtained by the three-channel UV spectrophotometer onboard Vega 2. Two “jets” were seen, one pointing toward the Sun, and another, stronger one nearly perpendicular to the sunward direction in the image plane. Since we refer here to this widely used term “jet” for the first time, let us make a brief comment. Apparently, it is used to describe a structured (linear or curved) brightness enhancement. However, as used in physical gas dynamics, a neutral gas jet is a region of collimated flow created by a localized source. It should not necessarily be a region of high brightness; conversely, regions of gas at rest (shocked gas or stagnation gas) are regions of high brightness, and not at all akin to jets. Furthermore, a gas jet is an isolated system, not the subsystem of anything; it cannot coexist at the same location with another, superimposed flow structure. But in the cometary case, the advocated “jets” result from image enhancement, and represent only a tiny fraction of the flow; they are flow details, not at all jets in the gas dynamic sense. This is not only a matter of terminology; it generates considerable confusion, prompting questions such as “What is the collimation mechanism?” Where there is no collimation, there can be no collimated gas. In the following we will continue to use the term “jets” in this manner, but the reader should be careful to remember that it designates only a localized detail in a large-scale gas or dust distribution.

Structures with a completely different appearance were discovered in P/Halley from ground-based observations of the CN distribution. These CN features could be observed during a much longer period of time than the previous gas jets, and exhibited definite periodicity:

1. Long-lasting spiral-shaped jets (A’Hearn et al., 1986), extending out to >6 × 10<sup>4</sup> km, were evidenced by azimuthal enhancement technique; Hoban et al. (1988) concluded from a dataset collected between April 9 and May 2, 1986, within 5 × 10<sup>4</sup> km from the nucleus, that the evolution of the morphology of the jets suggests the existence of a period of recurrence of 7.37 d. The current interpretation of these structures follows the original discussion of A’Hearn et al. (1986): The observed features are not due to local fluctuations in the density of a rotationally modulated fluid CN coma, but are due to parent species moving according to quite narrow spiral patterns (with thickness <3000 km at 30,000 km from the nucleus) and with a quite accurate radial velocity. “In the absence of a confining mechanism for the jets the inescapable conclusion is that the jets are composed of dust grains,” small enough to be invisible on white light images.

2. Ring-like, expanding “shells” were discovered by visual inspection of the CN images (Schlosser et al., 1986; Schulz and Schlosser, 1989). These shells (actually asymmetrical rings or haloes around the position of the nucleus) always dominated the outer coma, at a distance greater than 10<sup>5</sup> km from the nucleus, during the two-month observing period. Here again, the words “shell” or “ring” designate a localized enhancement of the brightness of a much wider distribution; furthermore, since it is observed in the emission of the nondominant CN molecule, it is uncertain whether it traces a local density enhancement of the whole coma, or a local variation in CN content. In any case, it would be incorrect to model it as an isolated distribution of gas: There can be only one global model of the coma, with the requirement that it reproduces the observed localized enhancement.

2.2. Comet P/Halley Gas “Jets” and “Shells”

Comet Hyakutake passed unusually close to the Earth on March 25, 1986. This led to the unique detection, during 10 successive nights, of spectacular arcs of OH, CN, and C<sub>2</sub>, centered on the antisunward cometocentric axis (at least in projection), having their apex on it at a variable distance (between 400 and 2000 km) and their convexity toward the Sun (reproduced in section 7.1). A detailed account of the observations is given in Harris et al. (1997) and in Rodionov et al. (1998). Both groups have interpreted the observations as the signature of an H<sub>2</sub>O arc, caused by some secondary H<sub>2</sub>O source. We compare and discuss the approaches in section 7.1.

2.4. Comet Hale-Bopp Spirals

Spectacular spirals appear in the large-scale images of Comet Hale-Bopp in March and April 1997 (i.e., shortly before and after its perihelion pass), for the observed molecules OH, CN, C<sub>2</sub>, C<sub>3</sub>, and NH (Lederer et al., 1997). These structures are conspicuous after image enhancement, but their real contrast must be high enough, as the authors state that they can be discerned even in the raw images. On
April 26, for instance, a single spiral is seen in the cleanest OH and CN images with an arm spacing of roughly $6 \times 10^4$ km. At an outer coma flow velocity of 1–2 km/s, this corresponds to a periodic modulation of the gas production rate with period 8–16 h. Lederer (2000) has interpreted these structures as follows. Parent molecules and fine organic dust grains are assumed to be ejected from the nucleus along straight lines, partly inside a certain number $N_j$ of cones rotating rigidly with the nucleus, partly inside the background space outside of the cones. The assumed nucleus rotation period is 11.3 h. Each cone is ascribed a relative production strength $S_j$, as is the background space, and a width $w_j$. The gas flux in any given dayside direction is assumed proportional to its $S_j$ and to the cosine of its angle to the solar direction. Once this is done, the daughter radicals are computed from a Monte Carlo procedure from the distribution of the primary molecules and organic dust, following Combi et al. (1993). A satisfactory fit is obtained assuming $N_j = 5$ cones; the cones and background are assumed to produce the same mixture of H$_2$O, HCN, and “parent-of-C$_2$” molecules, save one cone that, allowing for one-half of the OH in the spirals, is assumed to produce only H$_2$O and grains releasing only OH. The computed total OH productions from the “jets” represents less than one-half the total comet OH production.

Interferometric radiowave mapping of CO lines in the same comet on March 11, 1997, slightly before its perihelion (Henry et al., 2002), revealed time-dependent velocity-space structures in the emission. The measurements yield the line profile of the flux in each $1500$ km × $1500$ km element of the image. The intrinsic line profile of the CO line is negligible compared to the observed profile widths, so the latter are due to the Doppler shifts induced by the CO velocity distribution within the observed coma volume. Seen at a Sun-comet-observer angle of 45°, the profile usually exhibits two peaks approximately symmetrical with respect to the line center. The ratio of the number of molecules on both sides of the line center changes smoothly in about 7 h from =1.7 to =0.5, and back to nearly 1.5, suggesting a periodicity similar to that we have noticed above for the spirals (unfortunately, the CO observation duration is not long enough to confirm it). According to the observers, “this is indicative of a jet” rotating with the nucleus. The authors declare to have obtained good agreement with the data using “a 3-D model of the coma consisting of an isotropic contribution plus a conically spiralling jet of opening angle 30° and having outgassed 30% of the CO in the coma.”

The number and positions of the “active regions” postulated to explain the dust spirals, the OH, CN, and C$_2$ spirals, and the CO spirals are totally different, which is considered as “evidence for chemical heterogeneity in the nucleus” (Lederer and Campins, 2002).

After reading sections 5 and 6, the readers will have enough material at their disposal to form their opinion regarding the preceding interpretations. These interpretations are usually considered satisfactory because they “fit the observations.” This widely used argument is not sufficient. It is also required that the interpretation not be in conflict with any other related observation (cometary or not). The assumption that molecules move in straight lines conflicts with all estimates of the mean free path inside the coma, which is found to be smaller than 1 m near to the nucleus. Therefore, the inner coma should be modeled as a fluid, not as a set of non-mutually colliding particles. We return to this issue later in the chapter.

### 2.5. Dust Lightcurves

Usually dust coma isophotes are nearly perfect circles, and the radial slope of the coma brightness is well approximated by a power law vs. the coma radius with index $-1$. This led to the definition of the most frequently used tool to characterize the dust activity of a comet: the $A/\rho$ product (A’Hearn et al., 1995). In interpreting this product, two important assumptions are made: (1) the $1/\rho$ brightness variation is due to a trivial $1/\rho^2$ isotropic dust outflow; and (2) the dust grain outflow velocity is that applying to spherical grains ejected from a uniformly sunlit spherical nucleus — the “dust ejection velocity law.” We will comment on these assumptions in section 7.

### 2.6. Large-Scale Coma Dust Structures

In the following, we will focus on dust coma structures that can hopefully be related to the properties of the nucleus and of the near-nucleus gas-dust interaction. Presumably, the closer to the surface the structures are, the more suitable they will be for this purpose, as several effects perturb this relation; e.g., solar radiation reprocesses the structures because of the dependence, upon grain mass and grain composition, of the radiation pressure to solar gravity ratio $\beta$. Also, it is suggested from time to time that dust fragmentation is present in the coma.

Years of intensive ground-based dust coma image processing have shown that many dust comas that at a first glance appeared isotropic actually contain faint dust structures. Many of these structures strongly suggest the picture of a rotating inhomogeneous point source. The impression is even stronger if a motion picture of the observations is viewed — so strong that one is tempted to forget that these structures result from strong image structure enhancement, hence, as do the gas structures, characterize only a minor fraction of the whole coma brightness (however, in contrast with the gas, several different dust populations can evolve independently in the same region, since there are no dust-dust collisions). We have not found in the literature any estimate of even approximately how much nucleus dust production these figures represent. In fact, most authors process images by means of nonlinear algorithms, hereby losing any quantitative information on the extracted structures.

A good review of the classical interpretation of such structures was done by Sekanina (1991). A typical example of result is Sekanina’s (1981) model of the nucleus of Comet P/Swift-Tuttle. First, a trial-and-error procedure searches for recurring jet patterns in order to establish a nucleus rotation period. Then the time-dependent orienta-
tion of the most prominent jets is used to constrain the nucleus spin-axis orientation. As all dust of a given type is then assumed to move radially from the origin and with a common velocity $V$, an “activity map” is built in cometocentric longitude-latitude coordinates. This allows the identification of so-called “active” and “inactive” areas. The wider the jet, the more extended the active area is declared to be. Then, it is assumed that all grains in a jet were ejected at the same time, the curvature of the jet being due to the dispersion in $V$ and $\beta_s$ between the grains. Assuming that there exists a simple analytical expression $\beta_s = \beta_s(V)$ relating $V$ to $\beta_s$, the so-called “dust ejection velocity law,” the author derives, at each point of the jet axis, a value of $\beta_s$ and a value of $V$. The double jets that sometimes appear (e.g., in Swift-Tuttle) are interpreted in terms of two different dust chemical species: This shows the \textit{ad hoc} nature of the relation between active spots and jets.

Minor jets not used to constrain the nucleus spin state or not interpretable in terms of different dust chemistry are considered to be produced by \textit{ad hoc}-defined spots active only during a time segment adjusted to generate the observed jet. Thus, while the very use of the word “jet” suggests a long-lasting phenomenon, it is frequently assumed that the jet duration is short, sometimes lasting only minutes. Also, to avoid conflict between different observations, an active spot exposed to sunlight is often declared to have “deactivated” itself.

We do not question the fact that the “mechanism” thus offered, if carefully implemented, might give birth to the observed dust coma structures. But, in the literature, in particular in the two preceding references, it is used for reaching a much stronger conclusion: The “active area map” derived by this method is explicitly declared to represent the total activity map of the nucleus. In other words, (1) all the dust is assumed emitted from the active areas, (2) all the gas is assumed to originate from these active areas, and (3) the emission of the active areas is transient and, for some of them, occurs only one time. This analysis is often supplemented by a quantitative estimate of which fraction of the nucleus is active: For this evaluation, the active areas are assumed to consist of pure ice. Their total extent is found to be typically 5–10% of the nucleus surface.

Conclusion (1) meets with severe objections: Why would only a small fraction of the dust emitted by an active area go to the observed “jet,” while most of this dust would go to the background coma? And how are we sure that this is really the case? As no analysis is offered for the background coma dust (which makes up most of the dust emission), we are free to assume that it comes from the whole sunlit area of the nucleus. In fact, a much more natural assumption would indeed be that active areas (as derived above) are areas where, transiently, a tiny dust flux excess occurs, for any reason. In such a picture, all nuclei are essentially seen to emit dust homogeneously.

Conclusion (2) also meets with many strong objections. (a) As already pointed out by \textit{Whipple} (1982), it is arbitrary to postulate that a dust-gas mixture can be blown from discrete sources in a narrow collimated way; it is much more likely that the gas will diverge over a broad solid angle, and force the dust to do the same. On the contrary, if the gas emission is uniform (or nearly so) over the surface, its divergence should not be great (near-radial flow), hence trace variations in the dust content will be preserved outward. (b) By the same token, one does not see how a $\beta_s(V)$ relation derived for a strictly spherically symmetric gas flow would apply to a highly non-uniform and highly time-varying gas flow. (c) If, as likely, the mean direction of gas emission from a discrete area is set by the local topography (e.g., a pit), how is it that only sources active in precisely the postulated radial direction ever manifest themselves? (d) With a gas emission confined to transient localized areas, the torque applied by the reaction forces to the nucleus should be maximized, which renders the assumption of a simple, non-excited nucleus spin quite uncertain.

Finally, conclusion (3) also raises as many questions as it answers: What controls the switching on and off of an active area, if not the Sun?

In most papers, however, these questions are not addressed. In a few, they are raised, but not answered, or answered in a velikovskian way, suggesting, e.g., that the transient behavior is induced by the “opening and reclosing of cracks” in a surface mantle, and so on. But the most important point is that supporting quantitative physical modeling results are never presented, as if cometary activity stood outside the field of physical concepts and methods. This way of “explaining” the coma structure persists today, even though the first physical simulation of cometary activity casting a severe shadow on these explanations appeared 14 years ago (\textit{Kitamura}, 1990), and has been followed by the vast body of even more devastating gas dynamic results described in section 6.

2.7. Near-Nucleus Dust Structures

We use the term near-nucleus dust structures to refer to structures observed at a spatial resolution smaller than the nucleus size. Hence, data of this kind exist only for Comets P/Halley, P/Borrelly, and P/Wild 2. In the first case, the results are superbly described in the two-volume report published by the European Space Agency (\textit{Keller et al.}, 1995; \textit{Szego et al.}, 1995). In the two more recent cases, preliminary results have just appeared (\textit{Soderblom et al.}, 2004; \textit{Brownlee et al.}, 2004). In all three cases, the nucleus was also imaged. The spatial resolution was smaller (and position-dependent) for P/Halley, but the coverage was practically complete. For P/Borrelly, only part of the sunlit surface was imaged. The coverage seems to have been nearly complete for P/Wild 2.

The main result of the \textit{Giotto} flyby is a synthetic image obtained by the HMC camera, in which bright dust structures are seen attached to a restricted part of the nucleus edge (see Fig. 10 in section 7.2). This is typically described in the following terms: “... distinct jets emanated from active spots on the sunward side of the nucleus. Most of the elongated and structured nucleus appeared inactive” (\textit{Keller et al.}, 1994, p. 69). Figure 76 of the same reference
quantitatively reproduces the gross coma appearance, using a model described in Knollenberg et al. (1996): The dust distributions from three unequal circular active sources, placed on a sphere centered on the nucleus, are added. Each distribution is computed as if the source was isolated on that spherical nucleus. We return to this in section 7.2.

After enhancement of the azimuthal gradients of the HMC image, a wealth of fine radial structures appear, which the authors called “filaments” (Keller et al., 1994, pp. 83–85). A gas dynamical simulation of a process by which a narrow pencil of dust could be produced in an uniform ambient coma was developed (Knollenberg, 1994; Keller et al., 1994); an inactive circular area (100 m size) was assumed to exist as a defect inside a uniformly active surface; it was computed (not just stated) that a narrow pencil of dust is formed on its axis due to the convergence toward the axis of the surrounding gas and the resulting cross-axis motion of the dust. We return to this explanation in section 7.2. But we may immediately observe that this model result exactly supports Whipple’s (1982) criticism of the classical active area concept that we cited in section 2.6: A small-scale coma dust density maximum is found (not just assumed) to be due to a surface gas-dust production minimum.

Similar azimuthal enhancements were applied to the Vega 2 camera images. Here, due to a lower resolution, filaments could not be identified, but more than 10 directions of brightness enhancement were clearly distinguished (see pp. 208–228 of Szego et al., 1995). Enough view directions were available to conclude that “the jet sources formed a linear feature on the nucleus passing across the sub-solar point” (Szego et al., 1995, p. 72).

Both the Giotto and Vega observations are being reinterpreted by the global physical model described in Rodionov et al. (2002). We will discuss some of the results in section 7.2.

Finally, let us observe two essential differences between near-nuclear structures and distant coma structures:

1. The near-nucleus dust dynamics are controlled everywhere by the gas interaction, as already established by Whipple (1951). Hence it is unrealistic to claim to understand by means of visual observation the dust motion that is present before clearly stating how the (invisible) gas is considered to flow. In fact, when such images are presented to a gas dynamic scientist, the reaction is invariably that the observer is merely seeing plumes (i.e., dust in a gas flow).

2. Observation of the near-nucleus coma is concomitant with a determination of the nucleus shape. Therefore, realistic three-dimensional models can be developed. Furthermore, Comet P/Halley will return to perihelion in the year 2061, a horizon not totally discouraging for young scientists, and Comets P/Borrelly and P/Wild 2 will return much earlier. So, conclusions derived about these comets can be made under the form of precise predictions. As long as physical truth can result only from predictive-corrective iterations, these observations provide the first (and for the time being, the only) basis for a true physical study of comets. We return to these observations below.

3. PHYSICAL MODELS OF THE NUCLEUS-COMA INTERFACE

It is not possible to build a physical model of the coma without having a model of at least the outer layers of the nucleus, yielding at each point of the surface algorithms that allow computation of the temperature and gas and dust flux as a function of the solar direction (and distance). We say “algorithms,” because the coma conditions and near-surface nucleus conditions are mutually coupled and therefore must be computed self-consistently. One example of this coupling has long been recognized: The emitted dust can influence the visible and IR irradiation of the surface, hence react on the emission (Salo, 1988; Moreno et al., 2002). Another example, only recently documented, is that both net sublimation or net condensation are possible, even on the dayside surface, not only in the shadowed areas (Crifo and Rodionov, 2000) but on the sunlit portions as well (Crifo et al., 2003a). Hence, both nucleus interior models and coma models should be unbiased with respect to the value of the gas pressure at the surface, as well as with respect to the sign of the net surface gas flux.

3.1. Interface Description

At the present time, there are only indirect inferences about composition and structure of the nucleus. It seems that all authors follow Whipple (1950, 1951), who proposed that (1) it is a mixture of the ices of simple molecules and of refractory dust, probably with a complicated physical texture; (2) its outer, near-surface layers must be radially differentiated (the most volatile ices being absent from the outer layers). Estimating the stability of the various volatiles residing at or just below a nucleus surface, Whipple concluded that ices of the most volatile molecules (such as CO, CO₂, etc.) cannot survive one nucleus perihelion passage; this means that in comets approaching the Sun periodically, these volatile molecules must sublimate at some depth inside the nucleus and then diffuse toward the surface. It has been postulated sometimes that even H₂O ice itself could sublimate below a blanket of more or less cohesive dust. In all cases, the outer layers must be porous to permit gas effusion (actually, the voids created by the elimination of the volatile species already create porosity). The current speculations about comet formation in the early solar system also suggest that the nucleus as a whole may be a low-density, porous and brittle medium. One of the goals of the nucleus internal heat transfer models is to offer scenarios for this radial differentiation (see below).

Little consideration has been given in the cometary literature to the difficult problem of taking into account the surface topography. As already suggested by Whipple (1951), and confirmed by radar backscattering data (Harmon et al., 1999), the nucleus surface must be “extremely rough on scales of meter and larger.” The same must be true at smaller scales, due to dust ejection. Also, the surface is subjected to erosion — roughly 1 m will be lost per perihelion pas-
sage, which may imply locally stronger depletions. Thus the nucleus’ shape itself evolves both on a global and on a local scale. On a timescale of days, the submetric surface details will change. Keeping these facts in mind, the question arises of down to which level of accuracy does it make sense to describe these surface variations in the frame of a numerical model? The answer is difficult and depends upon the goal envisioned. The model of Rodionov et al. (2002) was constructed to handle the surface at a spatial resolution $\Delta = 50$ m, consistent with the Halley imaging data. The assessment of the Rosetta lander descent parameters requires a description of the surface such that short-term predictions of the near-surface coma structure (a few rotation periods) are possible. What this means in terms of spatial resolution of the surface has not yet been assessed.

Given that the nucleus rotates at an angular velocity $\Omega$, if the surface is to be described at the spatial resolution $\Delta$, for consistency this must be done with a time resolution $\Delta t = \Delta/\Omega$R. But should one use a true time-dependent model, or a succession of quasi-steady models?

As the near-surface nucleus material is potentially quite inhomogeneous in all respects (optical properties, volatility, porosity, granulometry, etc.), its relatively fast time-dependent illumination will induce both a horizontal and a vertical dispersion of the temperature(s). It is not at all proven that a quasi-steady-state temperature(s) distribution is ever achieved, nor even that local thermal equilibrium prevails. That is, there is no proof that, within a surface element $\Delta \times \Delta$, the various components (ice and minerals, for example) take on the same temperature. Actually this can only be expected if they are very intimately mixed and thermally well coupled. This problem has been considered to some extent by Kümle and Ulamec (1989), but certainly needs to be reinvestigated in a more general context. The same unanswered questions apply just below the surface.

The near-surface coma gas adjusts itself to a steady state extremely fast (typically within seconds), but this is not necessarily the case for the dust: Heavy grains can be accelerated to only fractions or small multiples of the nucleus escape velocity (meters per second) and hence they stay in the vicinity of the surface for times comparable to the rotation period. Thus, a thorough description of the gas production consists of a succession of steady-state maps of the mass density, velocity, temperature and of the various species’ mole fractions on a reference surface encircling the nucleus, at a spatial resolution $\Delta$ on the order of several mean free paths (m.f.p.), which is typically fractions of a meter to tens of meters — depending upon local solar zenith angle — near 1 AU from the Sun. For the dust production, not only the surface mass distribution $g_s(m)$ is needed, but the shape distribution $h_s$ (clearly out of reach) and a true time-dependent model may be needed as well for large grains.

3.2. Near-Surface Nucleus Interior

Volatiles residing below the surface can escape through pores and cracks in a rather direct way, if enough heat is transported to or created at the respective depth. There are in principle three ways in which the energy can be transported to a subsurface layer: (1) solid-state heat conduction via the solid matrix, composed of connected grains; (2) transport of heat by inward-flowing gases that recondense in the deeper/colder layers and release their latent heat there; and (3) penetration of the solar radiation into the ice and absorption in the interior instead of at the immediate surface; this can only happen if the ice is to a certain extent transparent and the radiation is trapped in the interior (solid-state greenhouse effect).

3.2.1. Porous ice models. The first thermal models of cometary nuclei (published in the 1980s and earlier) assumed nuclei to be nonporous ice/dust mixtures. Smoluchowski (1982) was the first to take into account porosity and gas flow through the pores. Subsequently, the heat and mass transport in porous, grainy ices was investigated in more detail by several groups (Squyres et al., 1985; Mekler et al., 1990; Steiner and Kömle, 1991a). The latter model was successfully applied to laboratory samples composed of artificially produced grainy ice (Kömle et al., 1991). The thermal evolution of larger ice/dust samples irradiated under space conditions was described by a similar model published by Benkhoff and Spohn (1991). The two latter models clearly showed that heat transport via the gas phase (energy transfer by sublimation/condensation processes) should play a significant role under “cometary” conditions. Otherwise it would be difficult to understand the measured temperature profiles. Another important aspect, studied by Kossacki et al. (1994), is the influence of grain sintering processes on the thermal evolution. Along the lines outlined by these models (which mostly included only water ice) multicomponent models were developed that allowed the prediction of the depths of various sublimation fronts as a function of the thermal history if a particular initial composition were given (Espinasse et al., 1991; Steiner and Kömle, 1991b, 1993; Benkhoff and Huebner, 1995; Kossacki et al., 1997). The current state of the art of these models is nicely described in the recent review by Capria (2002).

These models brought forward two important facts: (1) the possibly strongly reduced heat conduction caused by small grain contact area, as known from lunar regolith; and (2) the heat transported, by the migration of evaporated molecules along the thermal gradient, by sublimation and release of latent heat upon recondensation.

The basic equations describing this process are the conservation equations for energy (heat transfer equation) and mass (continuity equation)

\[
(1 - \psi) \rho g c_i \frac{\partial T}{\partial t} = \frac{\partial}{\partial x} \left( \lambda \frac{\partial T}{\partial x} \right) - c g \phi_k \frac{\partial T}{\partial x} - q H \tag{1}
\]

\[
\psi \frac{\partial \rho_k}{\partial t} = - \frac{\partial \phi_k}{\partial x} + q \tag{2}
\]

in which $\psi$ denotes porosity, $\rho$ specific mass, $c$ specific heat
content, λ heat conductivity coefficient, φ mass flux, H latent heat, q gas mass source term, and the subscripts i and g refer to the solid and to the gas phase, respectively.

While the basic equations used in these various modeling approaches are consistent, there are controversial approaches in the formulation of the boundary conditions: (1) In all the models noted above, free-molecular outflow into vacuum is assumed, i.e., the backflow of molecules from the coma toward the surface is neglected. (2) Solution of the continuity equation demands specification of the surface pressure and/or surface density of the emitted gas. Different authors use quite different boundary conditions here, from p = 0 at the surface to p = pS, the ice saturation pressure. Reduction of the two conservation equations (1) and (2) to one single energy equation is only possible if the surface pressure is assumed to be equal to the saturation pressure pS. This has been explicitly verified by Steiner et al. (1991). Something more realistic was used in the Espinasse et al. (1991) model. However, as a matter of fact, the highly nonequilibrated condition, in which the molecules are emitted, is not properly accounted for in any of the models. For a more detailed discussion of this point, see Skorov et al. (2001); this paper, as well as the previous one (Skorov et al., 1999), provided an important step beyond the previous modeling efforts, by presenting semianalytical solutions for the kinetic gas flow in tubes of finite length (assuming that the temperature profile along the length of the tube is known) and combining it with a numerical solution of the heat conduction equation. Thus it is not necessary to specify a separate pressure boundary condition at the surface, because the gas outflow from the tube is found directly from the temperature distribution along the ice tube walls and the associated local sublimation. The approach is similar to that already described in Kömle and Dettleff (1991), but with tubes instead of rectangular cracks.

3.2.2. Partially transparent porous ice models. A much more direct way to heat the subsurface layers of a comet nucleus exists if the ice is to a certain extent transparent. In this case the solar radiation is not fully absorbed or reflected at the surface, but penetrates icy layers down to a certain depth. It is then absorbed either by enclosed dust particles over a longer distance or by dusty layers with high optical thickness in the interior of the ice (see Plate 9).

The idea that the ice transparency could play an active role in the thermodynamics of comets and icy satellites is relatively old (Kömle et al., 1990; Brown and Matson, 1987; Matson and Brown, 1989). Recently it was reinvestigated by introducing the appropriate source terms into a more advanced thermal model (Davidsson and Skorov, 2002a,b) and applied to calculate the gas production rate of Comet P/Borrelly (Skorov et al., 2002) and to calculate the temperature profiles to be expected in artificial ice/dust samples (Kaufmann et al., 2002). The conclusions from these new calculations could modify the currently accepted view of cometary energy balance and gas production quite significantly. There are two important findings worth mentioning:

1. If forward-scattering dust particles are embedded in a transparent ice layer with a density realistic for comets, very dark surfaces (with a few percent albedo only, similar to that observed for Comet P/Halley in 1986) can be created. (2) With the same active area given, significantly lower gas production rates result, because the very surface is colder than in the case of “surface absorption” of the sunlight.

Laboratory experiments aimed at investigating this solid-state greenhouse effect more systematically and evaluating its significance for comets and other icy solar system bodies are currently under way (Kaufmann et al., 2002).

3.2.3. Comparative results. The vertical temperature gradient that develops in the near surface layer of the nucleus in response to solar irradiation strongly depends on the structure and composition of these layers:

1. For a compact, well-sintered dirty ice with little porosity the thermal conductivity is high, close to that of compact water ice. For long constant irradiation this leads to a temperature profile close to linear, but not steep, as calculated from classical models using the Klinger (1981) formula for the conductivity of water ice (a weak temperature dependence, compared to the exponential temperature dependence, which characterizes the heat transfer by gas sublimation/condensation).

2. If the ice is porous (open porosity) and composed of grains with low contact area, there will be a steep temperature gradient at the very surface (where the sublimating gas flows outward) and a rather flat temperature profile below, where the sublimated gas flows inward toward colder regions and recondenses there, because the effective thermal conductivity of the medium is high.

3. If the ice is covered by a loose dust mantle a few millimeters or centimeters thick, composed of refractory grains, a very steep temperature profile develops across this dust mantle, with a temperature drop of 100 K or more. This is due to the fact that the dust mantle in the low-pressure environment has an extremely low thermal conductivity, similar to that of lunar regolith. The conductivity of a dust mantle could be increased by orders of magnitude, if it contains organic components that might act as a glue between particles and cause some cohesiveness (Kömle et al., 1996).

4. If transparent ice exists as such, it may influence the activity of the surface in various ways. Depending on the gas permeability of the transparent layer, it may cause subsurface pressure buildup and violent activity if the gas pressure exceeds the tensile strength of the crust. A typical feature is the existence of a subsurface temperature maximum, as shown in the example below.

An example for the temperature profile that may develop inside transparent ice subject to the solid-state greenhouse effect is shown in Fig. 1. The main feature is that the maximum temperature occurs not at the surface, but a few centimeters below it. From there the heat is conducted away toward the surface. The position and sharpness of this subsurface temperature maximum depend on the absorption profile as well as on the thermophysical properties of the ice. Higher temperatures could be reached in the case of a
low gas permeability, possibly caused by the sintering and densification processes described before.

3.2.4. Stumbling blocks. The next logical step of the previously described interior models would be their combination with a near-surface coma model to derive the proper gas conditions at the surface. But here, a severe obstacle of most present nucleus models is that they are one-dimensional (and time dependent), and hence cannot accommodate the expected complex geometry of the real surface, nor even the smoothed approximation used in the coma models. It is unclear how this could be circumvented in the future.

Another consequential limitation of these models is that they assume that ice and dust share a common temperature at the surface (and in the near surface interior). This may be true for very fine dust, but cannot be true for large grains (pebbles, boulders). In fact, the laboratory simulations of KOSI have already found evidence of sizable dispersions in the surface temperature of samples of illuminated ice-dust mixtures even though only small amounts of fine dust were used in the experiment (in an unrealistic way) (Lorenz et al., 1995). In reality, surface dust is expected to be warmer than surface ice and will therefore radiate more thermal energy, hence its neglect overestimates the energy available for sublimation. The models with a surface crust are free from this criticism, but in general lead to rather small surface fluxes. In some models (e.g., Enzian et al., 1997, 1999), a first-order correction is made in the surface energy budget equation by the introduction of an “icy area fraction” f. We come back to this in section 3.3.2.

Let us notice that most nuclei size estimates assume not only that the surface is isothermal, but that it is really pure sublimating ice with no embedded dust — even in very dusty comets. For instance, the often quoted estimation “that 10% only of P/Halley’s surface is active” is based on such an unlikely assumption.

Efforts have been made to develop models capable of handling the geometrical complexity of real nuclei. Enzian et al. (1997) developed a “2-D 1/2 + t” approach for a spherical nucleus — in the sense that one-dimensional thermal equations (in the direction perpendicular to the surface) are solved at each point of the surface. Gutiérrez et al. (2000, 2001) extended this method to aspherical nuclei and Rodionov et al. (2002) developed a similar method for complex nuclei having spatial details of size comparable to that used in their coma model. However, for mathematical tractability, in all these models the nucleus interior physics has to be simplified to pure heat transfer.

3.3. Near-Surface Dusty Gas

The relevant modeling of the first tens of meters above the surface of an active nucleus is the only way to cross-correlate the thermal and dynamical evolution of the nucleus (mass loss, orbital perturbations, angular momentum changes) and the formation of its gas and dust coma. It is an unescapable task if one wants to interpret or forecast the coma structure, orbital evolution, or nucleus rotational state. Models that, for instance, prescribe “arbitrary boundary conditions” at the nucleus to compute the coma structure are intrinsically unable to provide any information on its evolution during the nucleus rotation (to interpret coma dust structures, for example.)

Unfortunately, the physical conditions in the near-surface gas-dust mixture can only be the subject of speculation: (1) the statistical properties of the surface topography and composition are unknown; (2) very little information is available about the dust; (3) where gas is released by the surface, it is not known whether it diffuses from below the surface, is sublimated from it, or both; (4) one is free to advocate more intricate processes such as dust fragmentation, sublimation from icy grains, etc. It is hard to believe, however, that tiny details of what happens here influence the structure of the observable coma, considering the existence of efficient smoothing processes inside the gas and dust coma: pressure for the gas; shape, mass, and compositional dispersion for the dust. Hence it can be hoped that effective (simplified) models of this region can represent in a roughly correct manner how the nucleus and coma are coupled. Of course, this can also be tested by developing alternative models that allow for different processes.

3.3.1. Near-surface conditions in a dust-free case. Elementary calculations show that a surface of pure ice submitted to solar illumination in free space assumes a temperature below the water triple point, and hence sublimates: It emits molecules according to a half-space centered Maxwellian distribution $M(0, T_{n})$, where $T_{n}$ is the surface temperature. The rate $Z^{*}$ (molecule/m² s) of the emitted molecules is, very roughly, $Z^{*} = c_{\odot} \cos \theta_{0} L_{0}^{2} / L_{0}^{2}$, where $c_{\odot}$ is the solar constant, $\theta_{0}$ the solar zenith angle, and $L_{0}$ the sublimation latent heat. If the gas diffuses from pores, the distri-
bution is no longer a half-space Maxwellian, but some other function defined by the pores’ geometry and temperature, and $Z^+$ is necessarily of a smaller magnitude than before. It is easy to compute that, near $r_h = 1$ AU, the mean free path of the molecules emitted from a given point against collisions with those from the adjacent points is very small (typically a fraction of a meter), hence a fluid atmosphere is formed. A flow pattern must establish itself inside this atmosphere, resulting from the complex surface distribution of $Z^+$. It is the purpose of the coma model to compute this flow. Here, two difficulties arise, first mentioned in the Russian literature (see references in Crifo, 1991a), and first discussed in detail in Crifo (1987).

The first difficulty is that the velocity distribution of the molecules returned to the surface is the downward part $M^+(v_0,T_0)$ of a Maxwellian function with some mean velocity $v_0$ and temperature $T_0$. Let us refer to $M^+(v_0,T_0)$ as the corresponding upward part, then the distribution of the emitted molecules must differ substantially from $M^+(v_0,T_0)$, otherwise the net gas flux at the surface would be vanishingly small. Therefore, in the immediate vicinity of the surface, the gas is not in an equilibrium regime (which requires a strict, or moderately distorted Maxwellian shape). This region must therefore be treated by gas kinetic methods (solving the Boltzmann equation, or BE). If this region is small, hereby defining a so-called surface boundary layer (BL), the much more efficient gas dynamic methods will be used to compute the flow outside of it (section 5).

The second difficulty is that in a gas, the flow regime depends upon the conditions holding at all boundaries of the flow. Therefore, the preceding $(v_0,T_0)$ near any point of the nucleus surface does not depend only upon the local values of $Z^+$ and $T_n$, but upon these values at all points of the surface. From the point of view of nuclear interior models, this means that the return flux $Z^-$ (recondensed onto the ice) is not proportional to the upward flux $Z^+$ (contrary to what is assumed in many recent cometary papers). In particular, there is no “simple” way of predicting whether $Z^-$ is smaller or greater than $Z^+$. This is not simply a cosmetic argument: Indeed, we shall see in section 6 that very plausible surface topographies lead to recondensation (instead of sublimation) over sizable fractions of the sunlit areas of the nucleus.

These two difficulties have been identified for quite some time in the rarefied gas dynamic literature, and continue to be the subject of advanced developments (see Cercignani, 2000, and references therein). The reader is urged to be cautious about the relevance of publications concerning these problems that do not refer to the aforementioned literature. The specific problems of integrating the BE over sublimating or condensing ice under simple geometries have long been solved exactly by analytic methods (e.g., Cercignani, 1981) or direct Monte Carlo simulations (DSMC) (e.g., Abramov, 1984). Notice in passing that for the reason given above, solutions can only exist for specific geometries. Of special interest is the plane-parallel solution, because any surface that is smooth enough can be approximated by locally plane-parallel elements. This is done systematically in all papers by Crifo and Rodionov cited hereafter. The accuracy of this approximation is found to be unexpectedly high (see section 6.3.3.).

In the plane-parallel method, the “surrounding flow” degenerates to the specification of one set $(p_0, v_0, T_0)$: the gas parameters at the distance where equilibrium flow is reached. This distance is found to be on the order of several tens of mean free paths. This implies that, for a comet near 1 AU, the thickness of the BL will be several meters to several tens of meters. Obviously, the mathematical smoothing of a real nucleus surface to $\Delta \approx 50$ m makes the plane-parallel approach usable, but it remains to be seen whether this smoothing is acceptable in itself. The computed structure of the BE is found to depend upon one free parameter, in accordance with the previously established fact that there is no way to predict $Z^-$ by consideration of the BL flow only. For this free parameter, the initial Mach number

$$M_0 = \frac{v_0}{\sqrt{k_B T_0/m}}$$

in which $k_B$ is Boltzmann’s constant, is usually used. It has been proven from first principles that no solution exists for $M_0 > 1$ (once more, contrary to some — inaccurate — results published in the recent cometary literature). Thus, one arrives at relations $p_0 = p_0(T_n,M_0)$ and $T_0 = T_0(T_n,M_0)$ at each point of the top of the BL, to be taken as the boundary conditions for computing the coma flow. Only when this flow is computed (see section 5) is $M_0$ known at each point, hence the return flux $Z^-$. This approach is fully described in Rodionov et al. (2002) and references therein.

3.3.2. Near-surface conditions in a dusty case. Perhaps pure ice volumes exist in some nuclei, but, in general, one expects the ice to be dirty. It follows that only a fraction $f = 1/(1 + \mathcal{R})$ (where $\mathcal{R}$ is the relative dust-to-ice mass content) of the surface is ice; the rest is dust (e.g., Crifo, 1997). As already stated, it is not expected that these two constituents assume a common temperature. Furthermore, dust, even porous, cannot produce vapor at the rate $Z^+$ of ice. The above picture must therefore be amended. An approximate modification of the above model has been proposed by Crifo and Rodionov (1997a), consisting in writing that the net upward flux is reduced (compared to pure ice) by the factor $f$. As already stated, this factor is used in many nucleus models, but unfortunately not in all.

Since the initial dust velocity is negligibly small, the fact that there is a large dust concentration inside the BL must still be taken into consideration. Based on the fact that the dust density is also large just outside the BL, and that, outside of it, exact dusty gas dynamic computations show that the perturbation of the gas is small, Crifo and Rodionov (1997a) assume that this is also true inside the BL. This analysis was done assuming a P/Halley-like dust mass distribution of spherical grains. It is possible that the situation could be different for other dust properties.
4. PHYSICAL MODELS OF THE NUCLEUS ROTATION

In the same way as it is impossible to build a coma structure model without matching it to a subsurface nucleus model, neither is it possible to reproduce most of the coma structures, whether directly observed or observed through lightcurves, without introducing a nucleus rotation model. This is true even of the near-nucleus structures, which critically depend upon the direction of solar illumination of the complex nucleus orography.

The nucleus rotation is governed by the equations for an asymmetric top (Landau and Lifshitz, 1976, section 36), taking into account the outgassing torque. Differences among researchers appear when evaluating the net recoil force \( \mathcal{F}_n \) and net torque \( \mathcal{T}_n \) exerted on a nucleus by its gas and dust emission. However, such computations are at the heart of the aerospace industry (rocket motors), so that reliable methods of computation exist in that field of study. For pure gas, both \( \mathcal{T}_n \) and \( \mathcal{F}_n \) can be computed by integration over any closed surface inside the gas flow, as long as mass, momentum, and energy (MME) are preserved. The best is to choose the top of the BL, where (1) the gas is in fluid regime (hence the integrals involve only \( p_0, V_0, T_0 \)) and (2) MME have been preserved as photochemistry and other effects are not yet at play. For dusty gas, the only exact method is the nucleus surface integration. However, for a P/Halley-like distribution, one can still integrate at the top of the BL, as stated above, because this kind of dust does not substantially perturb the flow inside the BL (in keeping with the fact that its global momentum is still negligible at this point).

An exact computation of \( \mathcal{T}_n \) and \( \mathcal{F}_n \) is possible only for nuclei with a known external shape and surface composition. The adoption of an arbitrary external shape when the real shape is unknown is unwarranted. The worst possible practice is to assume a spherical shape, since unrealistic symmetry cancellations occur during the surface integrations and since, in addition, the strict periodicity of the rotation induces a quite atypical quasisteady surface temperature distribution.

Rodionov et al. (2002) have computed \( \mathcal{T}_n \) and \( \mathcal{F}_n \) for P/Halley using the observed shape, a best-fit rotation mode, and assuming that the nucleus is uniform in composition. This revealed that both \( \mathcal{T}_n \) and \( \mathcal{F}_n \) vary considerably in magnitude and direction during the nucleus rotation. Belton et al. (1991) and Samarasinha and Belton (1995) calculated the nongravitational effects on P/Halley assuming that the dust jets seen far from the nucleus originate from active areas derived by tracing the jets back to the surface of an approximating ellipsoid. This is subject to two severe criticisms: (1) The correct shape should have been used, and (2) the definition of “surface jets” should have been done on the basis of fluid dynamics — this is the same criticism as when discussing Hale-Bopp spirals. We will return to the rotation of P/Halley in section 7.3.

Because \( \mathcal{T}_n \) scales as the third power of the characteristic nucleus size, while the inertia momenta scale as its fifth power, the angular acceleration due to outgassing scales as its inverse square. The effect of the torque is therefore expected to be maximum for very small nuclei. Indeed, Crifo et al. (2003b) have computed the torque and rotation of very small (subkilometric) irregular nuclei and found it to induce highly irregular, possibly chaotic rotational motions. On the other hand, the effect is modest for Halley-like comets, and must be negligible for very large comets (e.g., Hale-Bopp).

5. PHYSICAL MODELS OF THE COMA

By physical model, we mean an approach that tries to make full use of the latest available methods in applied physics. Unavoidably, only brief descriptions of these methods appear in the cometary literature. The reader unfamiliar with these methods will first have to get acquainted with the proceedings of the Rarefied Gas Dynamics Symposium held every two years. Examples of up-to-date textbooks covering most topics of interest here are Gombosi (1994) and Cercignani (2000). For information about fluid dynamics computational methods, see references in Rodionov et al. (2002), and for the Monte Carlo simulation methods, see Bird (1994).

Once methods are available, they must be applied. Here, the cometary medium must be described by selecting numerical values for all physical properties. This occasionally raises difficulties. For instance, collision cross sections between exotic molecules may not be known. However, the dominant molecules in the coma — \( \text{H}_2\text{O}, \text{CO}, \text{etc.} \) — have been the subject of in-depth studies in the laboratory and in the industry. Their physical properties (both in the gas phase and in the solid phase) are therefore available. Unfortunately, it is not uncommon to find works in which fancy numerical values are adopted for these properties. The nonexpert reader is therefore advised to check all values against such robust sources as the American Institute of Physics (AIP) Physics Handbook (last edition in 1972) or the Chemical Rubber Company (CRC) Press Handbook of Physics and Chemistry (84th edition in 2003).

While a gas molecule is a precisely defined entity, such is not the case for a “dust grain.” One usually ignores it by the “simplifying assumptions” that the grains are spherical, but this is unfortunate: Intuition as well as pioneering simulations using aspherical grains (Crifo and Rodionov, 1999) indicate that spherical grains have a totally atypical aerodynamical behavior: A flowing-by gas submits them only to a drag force, the lift and torque being null. It follows that all spherical grains of a given mass starting from a given point follow the same trajectory and have the same final velocity. Such is not the case for grains having the same mass, but varying shapes: Their trajectories differ, and their final velocities can be spread over orders of magnitude. In fact, even the initial orientation of the grain at the surface influences their trajectory and final velocity, with some
shapes and orientations preventing ejection, and others leading to high velocities.

Not only does it appear mathematically difficult to build a model of a near-nucleus coma with shape-dispersed dust, but it seems unlikely that one will ever access the needed input parameters of the model, such as the grain shape distribution and the distribution of initial orientation of these grains at the surface. From this point of view, the ultimate goal of cometary dust studies is to tell the truth rather than to lead to high velocities.

5.1. Gas Coma

The ultimate description of a coma is the set of distribution functions $f_i(t, \mathbf{r}, \mathbf{v}_i)$, defined as the number of particles of species $i$ (H$_2$O molecules, CO$^+$ ion, spherical olivine grain of radius $a_i$, etc.), having the cometocentric particle velocity vector $\mathbf{v}_i$ at the cometocentric position (vector) $\mathbf{r}$, per unit volume $d\mathbf{r}$ and unit velocity-space volume $d\mathbf{v}_i$. These functions are governed by coupled BEs (see, e.g., Gombosi, 1994; Cercignani, 2000). It is the goal of gas kinetic methods to solve these BEs. It can be done in exceptionally simple cases by numerical methods; otherwise, in principle, DSMC can be used (see Cercignani, 2000). However, DSMC requires forbidding computational resources as soon as interparticle collisions become important. Fortunately, in this case, solving the BEs is unnecessary because it is possible to predict the general form of the solution. For instance, in a gas mixture where near-thermal equilibrium prevails, the $f_i$ are exact or nearly exact Maxwellian functions of the gas mean mass density $\rho$, flow velocity $\mathbf{V}$, temperature $T$, and species mass concentrations $q_i$. It is therefore optimal to solve only fluid equations governing these quantities — the objective of gas dynamic methods. Computing the same flow by the two alternative methods, when possible, is beneficial, providing, in particular, cross-validation of the numerical methods.

It is a fairly difficult task to derive from first principles which method is best suited to deal with a rarefied flow such as the coma gas flow. The discussion is to be based on a comparison between the mean free path and the characteristic scale $L$ of the flow, as well as between the collision times and the timescales of the flow. This comparison must be done at each point. Additional limitations due to the mathematical methods also come into play; see a concise discussion in Rodionov et al. (2002), and more detailed developments in, e.g., Gombosi (1994) and Cercignani (2000).

To give an oversimplified summary, when the mean free path is much smaller than $L$, the gas is said to be “in fluid regime,” fluid equations apply exactly, and DSMC is useless; when the mean free path is much greater than $L$, the gas is said to be in “free-molecular” regime; in between, the gas is said to be in “transition regime.” In the two last cases, DSMC always applies, and fluid equations may or may not provide accurate results; the quality of their solutions must be evaluated case by case (see examples in Crifo et al., 2002a, 2003a).

5.1.1. Gas dynamic approach. The most general form used hitherto in coma studies is the single-fluid Navier-Stokes equations (NSE)

$$\frac{\partial}{\partial t} \rho + \nabla \cdot (\rho \mathbf{V}) = f_p$$

$$\frac{\partial}{\partial t} (\rho \mathbf{V}) + \nabla \cdot (\rho \mathbf{V} \mathbf{V}) + \nabla \rho - \nabla \cdot \mathbf{F} = f_v$$

$$\frac{\partial}{\partial t} (\rho h_0 - \rho) - \nabla \cdot \mathbf{q} + \nabla \cdot (\mathbf{f} \mathbf{V}) - \mathbf{F} \cdot \mathbf{V} = f_h$$

where $\rho$ is the total mass density, $h_0$ is the specific enthalpy, $p$ is the pressure, $\mathbf{q}$ is the heat conduction vector, $\mathbf{F}$ is the second-rank viscous stress tensor, and $\mathbf{F}$ is the total macroscopic force. The vector $\mathbf{q}$ can be expressed as a function of $T$, and $\mathbf{F}$ can be expressed as a function of $T$ and the partial derivatives of $\mathbf{V}$ with respect to the coordinates.

In $\mathbf{F}$ the radiation pressure force is generally negligible, and the nucleus gravity generally neglected, but there would be no problem with keeping the latter term in order to deal with Kuiper belt-sized objects if one wanted to. If the preceding equations are written in a non-Galilean frame, inertia forces must be included. For instance, if a nucleus-attached frame is used, it can be shown (Rodionov et al., 2002) that the Coriolis force $2 \mathbf{V} \times \Omega$ (where $\Omega$ is the nucleus angular velocity vector) is the dominant inertia force. It can also be shown that the ratio of the pressure force $|\mathbf{V}\mathbf{p}|$ to the Coriolis force decreases as $1/r$ with distance to the nucleus. In a very large comet with fluid region exceeding $10^4$ km (like Hale-Bopp), at such a distance a plausible $\Omega = 10^{-5}$ radian s$^{-1}$ sets the preceding ratio to about $1/10$: In other words, the Coriolis force is dominant — an hitherto unnoticed fact. Such may also be the case for less-productive comets if their nucleus has a higher spin rate. It implies that rotationally induced gas structures are to be expected at large distance.

The “source-sinks” terms $f_p$, $f_v$, $f_h$ at the r.h.s. of these equations allow for (1) the fact that the fluid interacts with photons, or with particles not belonging to the fluid itself (e.g., dust), and (2) possible inelastic processes that are internal to the fluid itself and that affect its momentum and energy budget. An important example of the first effect is photodissociation of H$_2$O, and an example of the second effect is partial H$_2$O recondensation into clusters (H$_2$O)$_n$, both yield a large $f_p$ term, a moderate $f_v$ term, and a negligible $f_h$ term. Another example is cooling through IR emission. It is not possible to incorporate dust inside the fluid described by the above NSE, because the huge mass difference between molecules and dust grains forbids the dust grains to share a common flow velocity with the gas (and even a common flow velocity between themselves). Furthermore, in typical coma conditions the grain-grain collisions are negligible, so that dust grains do not acquire thermal velocity spread or pressure. The dynamics of the grains must therefore be treated by separate equations (see below).
and their interaction with the gas must be represented by r.h.s. terms in the NSE. If the grains do not emit or condense gas, there is no $f_j$ term, and the two other terms have little effect on the gas, for a dust mass distribution of the kind found in Comet P/Halley (see Gombosi, 1986; Crifo, 1987). But if the dust is icy, it will condense or emit $\text{H}_2\text{O}$ and release or absorb latent heat, and this may result in strong perturbations of the gas flow (see Crifo, 1995).

The “source-sinks” terms, whatever their origin, must themselves be computed by solving so-called “rate equations,” to be solved simultaneously with the above ones. For a review of the forms of r.h.s. terms, see Crifo (1991a) and Rodionov et al. (2002).

For many applications, the Eulerian form of the equations (EE), simpler because it involves only first-order partial derivatives, can be used; this is obtained by setting $\xi = 0$ and $q = 0$. Formally speaking, the relative ranges of validity of the EE and NSE should be delineated using a dimensionless rarefaction parameter, the so-called generalized Knudsen number, defined in Crifo and Rodionov (1997a).

But comparisons with solutions obtained by DSMC show that the frontiers of these domains also depend upon the details of construction of the numerical method of solution (see Crifo et al., 2002a, 2003a). The existing results from such comparisons show that the NSE and even the EE provide acceptable solutions out practically the whole day-side coma of observable comets. Two restrictions are to be made, however: (1) The immediate vicinity of the nucleus surface must always be dealt with by gas kinetic methods. This is a very strong restriction, since only a correct treatment of this region warrants the obtention of a correct solution from the EE or NSE “downstream” in the coma. We have discussed this region in section 3.3.1. In all their studies, Crifo et al. treat this region by an algorithm based on the BE, hence they call their solutions “BE-EE” or “BE-NSE.”

(2) In the outer reaches of the coma, where dissociation products are dominant, it is presently not known to which accuracy these equations represent the real situation. This is because photodissociation creates the daughter molecules with high velocities relative to the parent velocity; for the fluid equations to be valid, it is necessary that slowing down of these products to the local velocity distribution occurs within one computational cell. This may not necessarily occur. Unfortunately, comparisons between NSE (or EE) and DSMC for such cases are not yet available.

Finally, let us comment about the presence of the time $t$ in the above equations. Given angular speeds $\Omega = O(3 \times 10^{-9})$ radian/s, near-nucleus gas speeds $V = O(300)$ m/s, and a modest lateral spatial resolution $\delta = O(0.01)$radian, one sees that steady-state gas solutions can be used out to distances $< (V/\Omega) \delta = O(1000)$ km. However, such time-stationary solutions can be obtained only by solving the time-dependent equations — with time-independent boundary conditions. The reason for this is associated with the fact that the gas flow is in (large) part supersonic; see references in Rodionov et al. (2002) for an extensive description of the methods of solution.

For modeling most observed coma gas structures involving distances much in excess of 1000 km, the time-dependent gas equations must be solved, but looking for a true time-dependent solution, not just for the limit for large times. To do this, instead of keeping the nucleus surface parameters constant, one must first compute a set of successive nucleus surface parameters, forming the boundary condition for the variable $t$. The solution is a set of successive three-dimensional coma structures. Obtaining it is an enormous undertaking in terms of computer resources. It has been achieved for the first time during the preparation of this text, and will be described only in forthcoming publications by Rodionov and Crifo.

However, observations involving distances in excess of 1000 km may be interpreted by a succession of steady-state solutions, if free-molecular conditions are reached before or near that distance: In such a case, the extrapolation of the gas parameters beyond 1000 km is trivial, and is not affected by photodissociation. This is, for instance, the case of the interesting Comet P/Schwassmann-Wachmann 1 discussed, e.g., in Crifo et al. (1999), which lends itself to velocity-resolved observations.

Finally, it may happen that the gas reaches, near 1000 km, transition regime conditions; in such a case, the validity of the use of fluid equations becomes uncertain, in particular due to the presence of photodissociation. It is then necessary to use a time-dependent DSMC, or to validate the use of fluid equations by comparison with DSMC results.

5.1.2. Direct Monte Carlo simulations. In a DSMC model, the evolution resulting from mutual collisions, of the individual velocity components, of the internal energy, and of the position coordinates of a large number of “weighted” molecules are monitored. Instead of introducing molecules at the rate $Z^+$, they are introduced at a somewhat reduced rate $Z/q$ ($q > 1$ can be position-dependent). The statistical consequence of the replacement of the extremely large number of real molecules by a much smaller number of simulated molecules is of course taken into account. Space is discretized into adjacent cells, and time into a succession of time steps. At each time, the number of mutual collisions of the molecules expected in each cell is evaluated, and the velocities and internal energies of a corresponding number of randomly selected pairs of molecules are changed using a binary collision model. During the next time interval, all molecules are moved. The procedure is reiterated until steady state is achieved. [For a description of the method, see Bird (1994); for an insight into the future of this method, see Bird (2001).]

The DSMC method offers specific advantages: (1) It is valid for any form of the gas velocity distribution function; and (2) the boundary conditions can always be formulated exactly (for instance, very complicated nucleus surfaces, on any linear scale, can be considered). But, as with any method, it can be (and is indeed sometimes) improperly implemented. For instance, three mandatory requirements are that (1) the chosen time steps must be much smaller than the mean collision time, (2) the typical cell dimension must
be much smaller than the local mean free path; and (3) the cell dimension must also be smaller than the characteristic flow scale \( L \). Unfortunately, information that allows the reader to check whether these conditions are satisfied is not always included in the publications. Finally, the method is computationally much less efficient than the solution of fluid equations, so it should not be considered to be a substitute for the latter, but should be used to complement them.

### 5.2. Dust Coma

If the dust perturbs the gas flow, its distribution must be computed self-consistently together with the gas equations. This was done in the old one-dimensional works reviewed, e.g., by Wallis (1982) and later on by Crifo (1991a), and in several of the two-dimensional works reviewed in Kömle (1990) and in the present section 6.2. In most of these works, all the dust mass loss was (unrealistically) assumed to be concentrated in single-size small spherical grains. This resulted in a strong perturbation of the gas. But with the mass being spread over a very large range of mass, as was found in Comet P/Halley, this effect disappears (see Gombosi, 1986; Crifo, 1987). While it is not possible to exclude very narrow size distributions, this presently seems to be permitted, hence one will currently solve first the gas equations and then the dust equations.

Conversely, as the gas density decreases outward, and the gas-dust interaction is a function of at least the square of their relative velocity — which also decreases outward — the acceleration and cooling of the dust by the gas stops at some distance from the nucleus surface (typically less than 100 km).

Since the “dust velocity” is typically one or several order(s) of magnitude smaller than the gas velocity, the range of validity of steady-state dust distributions is also an order of magnitude smaller, i.e., is typically only 100 km. Beyond it, standard interplanetary dust modeling techniques are to be used (see Fulle et al., 1999; Fulle, 2004). These methods are intrinsically time-dependent.

Here, we will only deal with the modeling of the region where the dust-gas interaction is sizable. If possible, it is appealing to use, for the dust, equations similar to those for the gas. However, the latter express the fact that mass, momentum, and energy are preserved during the many collisions occurring in each elementary volume. For cometary dust, collisions are practically absent. Hence the use of fluid equations seems unwarranted. However, it is correct to write that the mass, momentum, and energy of a set of co-moving particles are preserved. So, it is possible to group in the same “fluid” particles that follow everywhere the same trajectory. Since the aerodynamic acceleration depends upon mass and shape, these particles must have the same mass, shape, and initial orientation at the surface. This is still not sufficient: One must also make certain that trajectories do not cross one another at any point. This leads to subdividing the nucleus surface in areas such that dust grain trajectories emanating from any given subdivision never mutually intersect. We will present illustrative examples of this in sections 6.1 and 6.2.

In conclusion, it is possible to compute the dust distribution at small distances from the nucleus by a so-called “multifluid model” with possibly a very large number of fluids. Each fluid will be governed by “zero-temperature EE” with \( p = T = 0 \) and \( b_0 = c_s T_s \) (\( s \) is the grain type tag, \( c_s \) the specific heat, \( T_s \) the grain internal temperature). An other possibility would be to use a DSMC for the dust, seeding them inside the precedingly known gas solution (see section 6.2).

### 6. INNER COMA STRUCTURES AS REVEALED BY PHYSICAL MODELING

On the scale of the ground-based data spatial resolution (hundreds of kilometers at best), the zero-order coma gas flow is trivial: A point source placed in a near-vacuum can only provide a radially diverging flow; mass conservation produces \( a \propto (1/Vr^2) \) gas density decrease, the resulting pressure gradient accelerates the gas velocity, and kinetic energy conservation requires a concomitant gas temperature decrease; a classical analytic treatment demonstrates that the flow becomes rapidly supersonic (e.g., Wallis, 1982). The flow will ultimately “freeze” at some terminal velocity and temperature when collisions become rare (e.g., Cercignani, 2000, section 5.9). Innumerable illustrations of these effects are described in the gas dynamic literature. The solar wind is formed by a similar process (Parker, 1965). Of course, in the cometary case, there are also specific so-called “nonadiabatic” effects by which mass, momentum, and energy can be input or removed from the flow, thus altering its structure. For instance, a general gas dynamic theorem states that exothermic effects tend to render the flow sonic: If they occur in a subsonic flow, its Mach number is increased, and if in a supersonic one, the Mach number is decreased (possibly strongly enough to generate a shock).

Partial recondensation of H\(_2\)O molecules into molecular clusters and large amounts of fine dust may heat the gas. These effects are limited to the vicinity of the nucleus surface, because the first one is proportional to at least the square of the gas density, and the second one to the square of the gas density. Water photodissociation releases fast H and OH that, by thermalization, heat the gas; this occurs in a large part of the coma and limits the gas Mach number.

In the old cometary literature (1965–1990), the supersonic state of most of the coma and the preceding nonadiabatic effects have been recognized and studied in trivial one-dimensional geometry [see the review of Crifo (1991a), supplemented by Crifo (1993) and Crifo (1995) for posterior developments]. However, an essential implication of the supersonic state of the coma was universally overlooked in that literature: the tendency of the flow to form shock structures to adjust itself to nontrivial geometrical constraints, or in reaction to external perturbations. Gas shocks of inter-
est here can tentatively be divided in two groups: (1) “jet jet interactions” (a classical problem in the design of multiple thruster rockets); here, two gas flows meet — if one only is supersonic, it will form one steady shock; if both are supersonic, two shocks will be formed (e.g., Ni-Imi et al., 1992); and (2) internally generated flows; here, the heat deposition effect is strong enough to create a shock transition to subsonic state; in principle, H2O recondensation could create such shocks. Shocks are the canonical examples of structures in a nonturbulent fluid. Hence, it is somewhat surprising that the possible connection between shocks and coma structures was only first suggested by Kitamura (1990). However, geometrical constraints and strong non-adiabatic effects exist in the coma only at a short distance from the nucleus surface (abundantly illustrated below), a region not accessible to observations except in flyby and rendezvous missions. Even then, one essentially observes dust — not gas — structures, and dust cannot form shocks; it is surely not intuitive that near-nucleus dust structures trace gas structures, and an understanding requires advanced simulations of the kind described below.

All presently published physical model results refer to the near-nucleus coma. However, because of the scarcity of space missions to comets, most coma structures observed are located at very large distances from the nucleus. Are they just the result of the evolution to large distances of the near-nucleus structures, or do they result from other structuring mechanisms? We address this briefly in section 8.

6.1. Homogeneous, Spherical Nuclei

“Spherical nuclei” have been considered as the only paradigm for at least a half century of cometary speculations. In fact, isothermal spherical nuclei were assumed, even though no explicit mention of it was ever made. While this absolutely forbids any comparison with observational data, it is still a suitable paradigm to test new algorithms dedicated to a better physical representation of the coma (e.g., testing new radiative or chemical algorithms in the simplest possible way: one-dimensional equations). Here we will not deal with such uses; instead, see Crifo (1991a) and references therein.

On the other hand, it is not unreasonable for many purposes to consider that the outer coma is axially symmetric, as if produced by a sunlit, slowly rotating spherical nucleus. This has led to a number of two-dimensional spherical coma models, of variable merit, starting with Krasnobaev (1983) and continuing with Kitamura (1987), Kömle and Ip (1987a,b), Köröszmezey and Gombosi (1990), Knollenberg (1994), Combi (1996), Mueller (1999), Crifo and Rodionov (1997a, 1999, 2000), and Crifo et al. (2002a). In most works, EE were used, but in Kitamura (1987) and Crifo and Rodionov (2000), NSE equations were used as well. Combi (1996) used for the first time a DSMC approach. [The DSMC method discussed here should not be confused with the much less powerful “test particle Monte Carlo method” (MCTP) used previously by this author and a few others, and not discussed here.] Finally, the three methods were used and compared in Crifo et al. (2002a). In these numerous works, the boundary conditions at the nucleus differ: In all works except those by Crifo et al., the surface temperature and H2O flux are prescribed arbitrarily (most often, the surface flux is assumed to be \( \propto \cos z \theta \)); in the works by Crifo et al., either a CO flux is prescribed arbitrarily, or an H2O flux is derived from surface ice sublimation equations. The greatest difference in these input assumptions regards the nightside surface, which is either assumed to be inactive, or assumed to produce a uniform background flux of gas representing either a very small, or a sizable, fraction of the total dayside flux.

The computed structure of the dayside gas and dust coma is trivial (see Fig. 2) — notice, however, that a “gas velocity” or a “dust velocity” does not exist; both quantities are position-dependent. On the contrary, the nightside structure can be extremely complicated, as Fig. 2 indicates. It provides an ideal benchmark to discuss how the gas coma is formed, and with which accuracy it can be modeled. This led Crifo et al. (2000) to systematically simulate the range of possible nightside conditions by varying the nightside ice surface temperature, hence the nightside background gas production, using a heuristic parameter \( \kappa \ll 1 \): The thermal flux returned to surface elements in shadow is assumed to be \( \propto \cos z \theta \). Plate 10 shows the results. The \( \propto \cos z \theta \) variation of surface pressure on the dayside creates a lateral flow from noon to midnight; in the absence of night background (\( \kappa = 0 \)), the nightside surface is a cold trap for the gas. Therefore the flow from the dayside divides itself into one portion recondensing on the nightside and one portion escaping in the nightside hemisphere. The division occurs along a flow line terminating on the midnight axis at a point where the gas is at rest — a stagnation point. This gas at rest, as well as the gas reaccelerating upward and downward from it, form an obstacle for the arriving gas. Information cannot travel up a supersonic stream, so it cannot “guess” that there are obstacles ahead of it; it can only undergo a sudden transition to a subsonic state where information is received from the obstacle. Hence, as visible on the upper left panel of Plate 10, a conical shock is formed in between; it is, in fact, a converging-diverging shock (its diverging part appears conspicuously on Fig. 2). A very weak nightside surface background emission is enough to suppress the condensation; the emitted nightside gas is now an obstacle to the dayside gas, resulting in the formation of a weak converging shock attached to the terminator; inside the shock, the nightside gas accelerates rather slowly (Plate 10b). At a somewhat higher background level, fast acceleration of the nightside gas occurs (Plate 10c). The resulting supersonic stream interacts with the dayside one via a double shock structure; in between, sonic gas accelerates slowly outward (Plate 10c). Finally, a strong background creates the same kind of double-shock structure, but now the midnight stream stays supersonic all the way to infinity (Plate 10d).
Fig. 2. Coma around a sublimating homogeneous, spherical nucleus. The Sun is toward +X. (a) H₂O number density. (b) H₂O velocity. (c) 9-µm-radius dust grain number density. (d) Dust grain velocity. For these computations, the night background emission was assumed very low (κ = 0.01), so no nightside dust ejection occurred. (e) Trajectories of 2-µm-radius grains originating from various local times. One can see that, beyond the terminator, the trajectories mutually intersect. (f) Dust density. One can see that density peaks appear in the region of mutual trajectory crossings. On (e) and (f) (from Mueller, 1999), the night background is higher, and night dust emission occurs. Mueller (1999) calls the computed peaks “artificial,” which may be misleading (see text).
In real comets, a substantial night background production is expected from the nightside surface (for instance, from CO production, but other molecules are probable as well, e.g., CO₂, HCN, etc.). Therefore the flow structure should resemble that just described, save for the fact that the molecular composition will differ on the dayside and on the nightside.

The steep gas structures just evidenced (weak shocks) are found to translate themselves into dust structures, due to the fact that the dust particles are too heavy to accurately follow sudden changes in the gas direction: The trajectories of grains originating from both sides of the gas structures, and moving toward it, cross one another, creating a localized dust density enhancement. This is indeed found by most authors (when not found, inaccurate computational algorithms are to be blamed). For instance, Fig. 2e shows grain trajectories in the terminator shock region, where the gas velocity is about similar to that seen on Plate 10 (lower right panel). The sudden change in gas direction cannot be reproduced by the dust; instead, one sees dust trajectories mutually crossing and a resulting dust density enhancement (Fig. 2f). It is evident that on an image, these enhancements would be called “jets,” and one can see that such “jets” do not trace any dust grain trajectory, nor indicate the presence of any active area.

Further complication in the nightside coma structure would be introduced if one took into consideration a nucleus rotation: Coupling the gas dynamics to the nucleus interior heat transfer equations would introduce a surface temperature asymmetry, making the coma fully three-dimensional; in this sense, even the spherical, homogeneous nucleus still remains incompletely modeled today.

### 6.2. Inhomogeneous Spherical Nuclei

Many decades separated the first heuristic suggestions that dust structures could be due to nucleus “active areas” from the first gas dynamic simulations of the effect (Kitamura, 1990). Figure 3 shows most inhomogeneous spherical nuclei submitted to gas dynamic simulations to the present date.

The first computation was due to Kitamura (1986): One circular “active spot” defined by a Gaussian variation of the icy area fraction \( \alpha = g(z_0) \) is considered at a time when the Sun is on its axis. This assumption provides computational simplicity but not optimal significance. The author solved NSE equations for the gas. Plate 11a shows the gas distribution for a small spot surrounded by a strong uniform background: The formation of a conical weak shock (in reality a double-shock) appears clearly, matching the source flow to the background flow (and transforming the on-axis density maximum quickly into a minimum). Plate 11b shows what happens if dust is introduced: Dust density maxima are formed along the gas conical shock (on an image, the dust would appear as “emanating from two close active regions”).

Knollenberg (1994) revisited the same problem, but assumed another kind of background, \( \approx \cos z_0 \) on the dayside, and vanishing in the nightside. He solved EE equations. Plate 11b–d shows the result. One recognizes, on the nightside, the zero-background converging-diverging conical shock of Fig. 2 — the nightside is not sensitive to details of the dayside gas production. On the dayside, a conical (double) shock is clearly visible, similar to that in Kitamura (1986). When dust is introduced — on the dayside only, since there is no nightside background — it forms conspicuous enhancement in the vicinity of the gas shocks, for the reason already stated.

Kömle and Ip (1987a,b) considered a circular ring of increased activity, with Gaussian profile, superimposed on a two-step background (i.e., one dayside value and one nightside value), with the Sun placed on-axis. Unfortunately, we have verified by an unpublished recomputation that, as the authors suggest in their paper, their computational technique was inaccurate. Therefore, we prefer to comment here on the related (and highly accurate) results of Knollenberg (1994); see below.

The first paper to explicitly identify as shocks the gas density enhancements created by surface flux inhomogeneities was Kitamura (1990). Most importantly, this is also the first three-dimensional computation of a coma (EE were used). Two square spots are placed symmetrically about the noon axis on an inactive background (see Fig. 3a). Figure 4 shows the results, as recomputed by Crifo et al. (1995). The supersonic gas jets from the two sources interact strongly, forming a V-shaped double-shock structure (in the symme-
Comets II

Outside the symmetry plane, the double-shock structure is quite complicated in shape, owing particularly to the square shape of the sources (A. V. Rodionov, unpublished data, 1998). This is very similar to what is observed in multiple thruster rocket motors (e.g., Ni-Imi et al., 1992).

If dust is introduced in the flow, the result can be seen in Fig. 4d. One can see, once more, that neither the two density pencils at the edge of the central dust density enhancement, nor the enhancement itself, project themselves to the active spots. The origin of the central enhancement follows from the fact that the dust emitted by the left spot covers the angular sector 75°–130°, while that emitted from the right spot the sector <50°–105°; this means that in the region 75°–105° at each point there are two different directions of dust motion. In other words, dust density maxima do not reveal dust trajectories; furthermore, there is nothing like a “dust velocity direction” since dust grains moving along distinct directions coexist. Mathematically speaking, two different sets of governing flow equations must be solved — a so-called “two-fluid model.” This was done in Crifo et al. (1995) but not in Kitamura (1990). The result is that the dust density distribution in Kitamura (1990) is not accurately computed in the interaction region (Fig. 4b). However, the position of the density enhancement is correct. In more complicated geometries, in the interaction regions many different directions of dust motion may exist, requiring the use of many dust fluids. (This requirement should not be confused with that of also using different dust fluids if different kinds of dust grains are considered.) This quickly becomes unmanageable, and one has to content oneself with the indicative one-fluid (per dust type) method, or use a DSMC.

Knollenberg (1994) considered a problem resembling that of Kömle and Ip (1987a,b): An active circular spot is placed on an inactive surface (no background; see Fig. 3b). His computed gas distribution is shown on Plate 12. It was later recomputed by Rodionov and Crifo with strictly identical (unpublished) results, which we use to display additional details of the solution. The supersonic gas converging from the ring to the axis “interacts with itself,” forming a diverging conical weak shock with apex at a small distance over the surface. Below this apex a complicated low-velocity (subsonic) region is formed close to the symmetry axis over the inactive surface; it includes a circular stagnation line, parallel to the surface and centered on the axis, around which the gas whirls (the flowlines form a torus). Inside the conical shock, the slow gas is reaccelerated, forming a “daughter jet.”

Fig. 4. Spherical nucleus with two identical active areas (Kitamura, 1990; Crifo et al., 1995). (a) Isocontours of the gas density; (b) isocontours of 0.1-µm-radius grain dust density, computed from single-dust-fluid equations. (c) Same, on an enlarged geometrical scale; (d) same, computed from two-dust-fluids equations.
It is to be expected with such a gas flow structure that all dust trajectories starting from the inner edge of the active ring are directed toward the axis and mutually intersect there. The vicinity of this axis is thus expected to be a region of enhanced density. This is indeed what Knollenberg (1994) finds using (appropriately, but without saying it) a Monte Carlo simulation (Plate 12b). This is probably the most spectacular confirmation of the basic fact — already hinted at by Whipple (1982) — that the vertical of an inactive spot must be a dust-density maximum.

It should be noted that the symmetry that is present in all the preceding solutions is quite artificial: It follows not only from the assumption of a very simple nucleus geometry, but also from the equally strong assumptions of (1) solar illumination along the symmetry axis or (2) absence of illumination control of the gas production. The first assumption is not possible because of the nucleus rotation, and the second would require day-night symmetry in the appearance of the coma.

The results remains that (1) small inactive areas create usually coma dust density peaks, not minima, and (2) dust-density maxima do not trace only surface production maxima, but gas shocks as well.

Crifo et al. (1995) added to their duplication of Kitamura’s (1990) work the case of three aligned identical sources, with the Sun on the symmetry axis. In such a case, two “secondary” gas jets are formed in between the sources, and these two gas jets themselves interact at a greater altitude, to create “second-generation” weak gas shocks. The dust is not sensitive to these second-generation structures (at least at the moderate production rates considered) because gas-dust uncoupling has already occurred. Crifo and Rodionov (1997a) use the four equal rectangular active areas shown on Fig. 3d for several directions of solar illumination. The results, of course, reveal the formation of shocks similar to those discussed above, and illustrate for the first time (owing to the three-dimensional capability of the numerical code) the deformation of the near-nucleus coma with changing solar direction.

To conclude this quick overview, let us address the already mentioned configuration proposed in Knollenberg et al. (1996) to account for the gross Halley dust coma appearance during the 1986 Giotto flyby (Fig. 3c). These authors compute the gas and dust distribution from the three proposed unequal circular active areas as if they were alone with the Sun on-axis, and then add up the resulting dust densities. Instead, Plate 13 shows the gas flow correctly computed from solving in three dimensions the EE, and the dust flow computed from a four-fluid model — one fluid for each area, plus one for the background (Crifo and Rodionov, 1998). One sees that the results do not resemble three similar structures. First, the difference in solar zenith angle between the areas results in strong deviations between the gas (and dust) outflow patterns from one another, and from that of an isolated on-axis illuminated jet; second, the gas jets from the three areas interact, forming weak gas shocks and secondary dust-density maxima. But, as we shall see, there is another, fundamental reason why such a model cannot account for Halley’s coma: Halley’s nucleus is anything but spherical, and the outflow from a sphere cannot be “pasted by hand” on any aspherical body.

6.3. Homogeneous, Aspherical Nuclei

It is evident that cometary nuclei cannot be spherical, hence the investigation of aspherical nuclei is the central requirement of cometary activity models. This raises immediately the question of down to which scale one wants — or can — simulate a nucleus. Figure 5 presents examples of shapes that have been subject to investigation to date.

Fig. 5. Examples of aspherical homogeneous nuclei treated in the literature. (a) Triaxial ellipsoid (Crifo et al., 1999; Crifo and Rodionov, 2000); (b) bean-shaped nucleus (Crifo et al., 1997b, 1999); (c) apple-shaped nucleus (Crifo et al., 2003a); (d) top-shaped nucleus (Crifo et al., 2003a); (e) Muinonen shape (Crifo et al., 2003b); (f) Halley nucleus (Crifo et al., 2002b; Rodionov et al., 2002).
[To this one should add that Rodionov et al. (2002) have also studied about 17 different shapes derived from the Halley shape shown in Fig. 5 by additional and more-severe spatial filtering, and that several variants of the so-called “Muinonen shapes” have been studied in Crifo et al. (2003b).] One sees that at the present time only relatively smooth shapes have been considered. This is due in part to mathematical limits, but also partly because the best Halley nucleus image resolution was only $\approx 50$ m. The recent images of part of the surface of P/Borrelly (and of most of the surface of P/Wild 2) reveal that the surface is rich in very small topographic features. It may be expected that the fine gas structures expected from fine topographic details will be quickly filtered out by collisional smoothing, making the use of a smoothed surface suitable for modeling an imaged coma, given that line-of-sight integration should also smooth out such details. For the in situ sampled data, however, it is evident that surface modeling down to the metric scale will be required.

Not only are real nuclei aspherical, but they cannot be uniformly convex, as with all other small bodies of the solar system. Common sense indicates that the flow of gas over or near a totally convex object must be much simpler that that around or near an object with concavities. Hence, maximum attention must be paid to the latter.

6.3.1. Triaxial ellipsoid. A nonspherical, but still convex, nucleus can be considered to be merely a variant of the spherical nucleus: Figure 6 shows the case of a triaxial ellipsoid assumed either to outgas CO in a nearly uniform way, or to produce $\text{H}_2\text{O}$ through surface sublimation [see a detailed description of the latter in Crifo and Rodionov (2000)]. In the first case, one observes density kinks at the

---

**Fig. 6.** The coma around an homogeneous triaxial ellipsoidal nucleus. (a) Gas temperature at heliocentric distance of 3 AU. (b) Gas temperature at heliocentric distance of 1 AU. (c) Gas number density at 3 AU; (d) gas number density at 1 AU. At heliocentric distance of 3 AU, it is assumed that CO diffuses out with a nearly uniform surface flux; at heliocentric distance of 1 AU, it is assumed that $\text{H}_2\text{O}$ is sublimated, the Sun being in-plane in the direction of the angular graduation 45°; the nightside background parameter is $\kappa = 0.0275$. 
tips of the ellipsoid, and relatively smooth bulges over the flatter sides of the nucleus. This is in accordance with the fact that the local surface radius of curvatures scales the local near-surface flow, as explained in Crifo (1991a). Of course, such kinks will lead to the formation of dust density enhancements owing to the trajectory-crossing effect described previously. In the second case (solar-driven sublimation), the day-to-night pressure difference is enough to blow away the kinks at the tips, at least for the adopted direction of illumination. As for the general coma layout, let us observe what turns out (from many yet unpublished computations) to be a general behavior: (1) the nightside coma is structured around the antisolar axis, and has the same complicated patterns as in the spherical case, but distorted; (2) the dayside coma is not structured around the solar direction, but more around the normal to the flattest area of the surface (minimal curvature), again a consequence of the local curvature scaling just mentioned.

We may expect that, during the nucleus rotation, in the case of CO diffusion a nearly invariant density pattern co-rotates, whereas in the case of H$_2$O sublimation, the day-to-night density asymmetry will persist, although modulated by rotation-induced deformations (we return to this in section 6.5). Note that in the CO case there is, however, a significant day-night asymmetry in gas temperature (due to the corresponding surface temperature asymmetry), and therefore a significant day-to-night velocity asymmetry (not shown). In the H$_2$O case, the temperature (hence velocity asymmetry are quite small, due to the temperature buffering effect of the sublimation.

6.3.2. Bean-shaped nucleus. As stated earlier, the smooth gas flow possible around smooth, convex nuclei is impossible around realistic objects, because they have concavities; furthermore, in the solar-driven sublimation case, the pattern of shadows associated with the concavities subdivides the gas production into “effective” discrete active areas. For these two reasons, even a homogeneous nonconvex nucleus is expected to produce a coma structured by weak gas shocks. The pattern of these shocks will change during the rotation, following changes in the dayside shadow pattern. With this in mind, Crifo and Rodionov (1997b), for the first computation of a nonconvex nucleus, designed the “bean” shape shown on Fig. 5b. This shape has two planes of symmetry, hence an axis of symmetry. [This shape was also used for two-dimensional + 1/2 heat transfer computations by Guttiérrez et al. (2000),] Figure 7 shows the gas and dust-number density in three mutually perpendicular planes across the nucleus, under solar illumination from a direction inside the main symmetry plane. The dust was computed from a two-fluid-per-size algorithm.

There are two maxima of gas production around the two subsolar points, and the two associated flows from their vicinity interact to form two shock structures (of roughly hyperbolic shape). For the selected orientation, there is not yet partial shadow inside the cavity, but the left flank has less inclined illumination than the right flank, so there is an overall pressure gradient toward the Sun, which induces an inclination of the shock structure toward the Sun. The associated enhanced dust-density region is formed, as in the Kitamura (1990) case, by overlap between dust from the two subsolar areas. We will return to this coma below when discussing the effect of the nucleus rotation.

6.3.3. Apple-shaped nucleus. All preceding results, whether referring to comets near 1 AU from the Sun, or to larger distances (Crifo and Rodionov, 1997a,b, 1999), were obtained from EE equations insensitive to the absolute gas density. However, fluid interactions disappear at vanishing gas densities. Crifo et al. (2002a, 2003a) have started to investigated down to which low levels of gas production the weak shocks computed from EE are real. The method used is to compare NSE results with DSMC results. For this purpose, shapes somewhat simpler than the bean were considered, e.g., “apples” and “tops” (cf. Fig. 5). The surprise was that new effects were discovered during this work — and now it is not known up to which high level of gas production they persist!

Figure 8 shows the flow inside the sunlit cavity of an apple-shaped homogeneous icy nucleus with characteristic size $\approx 15$ km and assumed near the orbit of Jupiter (its computed total H$_2$O production rate due to sublimation is $Q = 3.3 \times 10^{26}$ molecules/s). One notices first that this flow is separated from the external flow (relative to the cavity) by the separating streamline eB, where B is a stagnation point on the axis, and e is on the surface; this means that no gas from the cavity escapes to free space! The flow inside the cavity divides itself into two closed cells: one is the triangular cell aAc, where a is the bottom of the cavity, A a second stagnation point on the axis, and c a point on the surface (see Fig. 8c); the second cell is the rectangular cell cABe (see Fig. 8a). In the first cell, the gas emitted from the segment ab (representing 0.06% of $Q$) is recondensed on the segment bc, without acceleration to supersonic state. In the second cell, the gas emitted from de (representing 0.2% of $Q$) is recondensed on dc, in part supersonically. Notice that the two flows inside the two cells rotate in opposite directions, as in convection cells. One additional feature of this fantastic flow structure is the presence of three sonic lines (SL1, SL2, SL3) transverse to the symmetry axis. Note also the broad size of the subsonic region (shaded on Fig. 8). Thus we see that, even though the cavity surface is fully illuminated by the Sun, there exists a band db where it condenses the ambient gas; this trapping effect explains why the cavity flow is confined inside the cavity, in spite of its wide opening. Such a flow trapping was not observed in the existing “bean” simulations, probably because this cavity has the shape of a saddle, i.e., is much more widely open to free space. On the other hand, it is not yet known whether the effect disappears in the “apple” at small heliocentric distances, and appears in the “bean” at very large distances.

Note that in the inner cell the Knudsen number $Kn = 1$ (cf. Fig. 8d); also, our approximating BE-NSE approach to handle the near-surface conditions, based on nearly plane-parallel sublimation or condensation, is here in principle
Fig. 7. Structure of the gas and dust coma around a bean-shaped nucleus (Crifo and Rodionov, 1999). The top panels show the main symmetry plane XOY; the Sun direction is in this plane, at the graduation 45°. Bottom panels show the secondary symmetry plane YOZ and plane XOZ. The lefthand panels show the decimal logarithm of the gas number density; righthand panels show the decimal logarithm of the dust number density (9.11-µm-radius grains).
invalidated by the fact that the two effects coexist at neighboring points. Even so, an impressive good agreement between the BE-NSE and DSMC solutions is obtained.

Finally, as to the disappearance of weak shocks at very small production rates, it was observed in the simulations, but this does not mean that structures in the gas density disappeared: Even in strict free-molecular outflow the gas density is quite uneven, hence dust small enough to be dragged away will exhibit sharp distribution structures.

The dust distribution created by the apple-shaped nucleus for on-axis illumination of its cavity resembles very much that obtained from the bean shape in its main symmetry plane (Fig. 4 of Crifo et al., 1997b), with one notable difference: Because the “apple” nucleus is rotationally symmetric and the Sun on-axis, so is the dust density, which thus is obtained by rotating the symmetry plane distribution around the symmetry axis; a narrow on-axis centered conical pencil of dust is formed (instead of two quasihyperbolic surfaces for a “bean”). Of course, this simplified geometry is broken as soon as the Sun leaves that axis.

6.3.4. Top-shaped nucleus. Figure 9 shows the computed gas and dust distributions when the Sun is on the symmetry axis of a top-shaped homogeneous icy nucleus. The presence of a partially shadowed cavity creates two separate gas jets whose interaction produces weak shock surfaces (one crosses the circular graduation near 70°, the other near 20°). Then, the dust is thrown into a conical pseudojet in the vicinity of this interaction. The conical
geometry will be broken for off-axis illumination. The dust density shown here was computed from a single-dust fluid approximation, and therefore is only indicative of the correct distribution, as discussed previously.

6.3.5. Realistic shapes. While consideration of simple shapes are a must for demonstrating the basic physical processes at play in the near-nucleus coma, realistic comas can only be produced by consideration of plausible nucleus shapes. Unfortunately, there exists no bank of cometary nucleus shapes from which the expression “realistic” could be defined, since only P/Halley’s nucleus shape has been determined (Keller et al., 1995; Szego et al., 1995, and references therein). Part of P/Borrelly’s nucleus has been recently imaged, but the whole shape is not known, hence the computation of the gas flow even assumed to emanate only from that part is impossible. It is hoped that the recently acquired images of P/Wild 2 will provide the second full nucleus shape. At any rate, within the frame of the Rosetta mission definition studies, it is necessary to have a set of “plausible comas” at hand. For that purpose, Crifo et al. (2003b) have considered nuclei having random shapes with the same statistical topographic properties as those of the imaged asteroids, following Muinonen (1998) and Muinonen and Lagerros (1998). An example of one such shape is shown in Fig. 5e. Such shapes were also used for thermal modeling in Gutierrez et al. (2001). The computed comas are complicated, as is the topography. We will not discuss them here, but will focus on the equally complicated case of P/Halley’s nucleus, discussed in sections 6.4 and 7.2.

6.4. Inhomogeneous, Aspherical Nuclei

A real nucleus can be expected to be both complicated in shape and inhomogeneous in composition. Since both effects lead to a structured coma, one does not see on which basis the observation of structures should be automatically attributed to compositional homogeneity. The question in fact arises whether these two effects contribute equally to structuring the coma, or whether one is dominant. There is at the present time no universal answer to such a question, in view of its extreme difficulty. Only the case of Halley’s nucleus has been studied (Crifo et al., 2002b; Szego et al., 2002). The complexity of the gas coma, even assuming Halley’s nucleus to be homogeneous, can be judged on the two left gas density panels (b,d) of Plate 14, corresponding to two different image planes and solar illumination directions. Plates 14b,c show the same information if the nucleus is assumed to have the random distribution of Gaussian circular active areas of average size comparable to the size of the topography details, shown on Plate 14a. Unexpectedly, the differences are very minor. In particular, most weak shocks are present in the two cases, at about the same location. Of course, it is premature to generalize this result to all conceivable nucleus shapes, inhomogeneity patterns, and solar illuminations. Yet we believe that this result is extremely instructive. It suggests that, at least for P/Halley’s nucleus, only extreme assumptions with respect to the inhomogeneity could significantly manifest themselves in the coma structure. Even though this was not yet computed, it is unlikely that the three-active regions pattern of Plate 11 placed in any manner on Halley’s nucleus will produced anything like the sum of three isolated cylindrical jets. We will see in section 7.2 how the results of Plate 14 compare with the observations.

6.5. Influence of the Rotation on the Coma Structure

Plate 15 shows the deformation of the near-nucleus dust coma around a bean-shaped nucleus with a simple rotational motion (no precession or nutation). One sees that the V-shaped structures rotates, but at a slower rate than the Sun, and with deformation, until disappearance for part of the rotation. This is an extreme example, however; for a case where the Sun would rotate on a cone with axis tilted to the “bean” axis, one can anticipate (based on the result showed
here) that the V-shape structure will rotate nonuniformly, with deformation, but without disappearing. The important point is that the apparent axis of rotation will not be that of the nucleus, and that its apparent angular velocity will not be its real angular velocity.

It is important to notice that the behavior of the near-nucleus coma just shown differs from that of the distant coma: The latter is a consequence of the evolution of the former during one or several nucleus rotation(s). For the dust, the physical process is the standard collisionless effusion in the solar gravity field reduced by radiation pressure. Hence, from the results shown on Plate 15 one can independently compute the outer coma dust distribution at any point and time. Such is not the case for the gas coma, as its collective behavior usually extends to very large distances; one can only compute at once the time-dependent structure of the whole coma. We return to this in section 8.

6.6. Conclusion: Formation of the Near-Nucleus Coma

The preceding model results are significant enough to draw many robust conclusions concerning the formation of the near-nucleus coma. Even though this region is rarely observable, it is clearly impossible to accept any paradigm concerning the currently observed outer coma that would violate any of these conclusions, which we summarize now:

1. The essential result is that the near-coma gas outflow is extremely sensitive to the surface topography. Roughly speaking, “hilltops” and high-\( f \) areas create supersonic outflows, “valleys” and low-\( f \) regions bounded by high-\( f \) ones create subsonic outflows, but simple rules to “guess” more precisely what the gas flow pattern is do not seem to exist — fluid dynamics cannot be guessed. Hence it would be physical nonsense to attribute near-nucleus gas structures (hitherto unobservable) to surface inhomogeneity only.

2. The gas outflow is a global property; there does not exist any physical mechanism by which some fixed structure (e.g., a conical jet) could be created at any given place, independently from the surrounding environment. Instead of coexisting without changing when “placed” on a surface, the structures created in insulation will unescapably be modified when placed in a surrounding environment, be it a background emission or some other similar structure. This is independent of how the gas is produced (surface sublimation or subsurface diffusion) and independent of its chemical composition and production rate. The gas interaction by definition does not stop at a small distance from the surface. Hence the paradigm of several simple, noninteracting structures advocated in most heuristic models of the outer coma described previously is physically impossible.

3. The assumption of one (or even more so, several) rotationally-invariant gas jet structure(s) modulated only in magnitude by a cos \( z_0 \) factor (found in many heuristic interpretations) is physically impossible; if there is such a solar modulation of the surface gas flux (as is the case for solar-driven gas production), then the “jet structure” will change its appearance (as shown in Plate 15). If one insists that the observations require a rigidly rotating gas jet, then the gas production should not be due to solar control, and therefore there should be no difference in observed activity between the day and nightsides of the nucleus, contrary to observations.

4. Inner coma dust structures are always associated with the gas structures of the inner coma. The formation of additional near-nucleus structures due to fluctuations of the dust concentration in the ice (i.e., a varying dust load in the gas flow) is possible, but, to interpret observed near nucleus dust structures, one must first determine which structures trace gas structures, and only when they have been identified can the remaining ones be attributed to dust-load inhomogeneities.

5. Dust structures resulting from gas structures do not trace dust grain motions; quite the contrary, they trace dust trajectory intersections. Their projection to the surface does not bear any simple relationship to the pattern of dust production at the surface.

6. Tracing dust grains back to where they left the surface across the gas-dust interaction region raises extreme difficulties: One would need to know not only their mass, but their shape as well, and even their exact orientation in the coma.

7. The impossibility of rotationally invariant near-nucleus gas structures implies a similar impossibility for the near-nucleus dust structures.

We will return in section 8 to the implications of these conclusions on the formation of outer coma (i.e., very large scale) structures. But first, let us see how the inner coma model results compare to observational data.

7. MATCHING PHYSICAL MODELS TO OBSERVATIONAL DATA

7.1. Hyakutake Arcs

Figure 10 shows an example of the nightside coma gas arcs observed in Comet Hyakutake. The arcs were also detected in OH by Harris et al. (1997), but there was no associated dust feature. Both Harris et al. (1997) and Rodionov et al. (1998) postulated that the arcs could only be the signature of an unobservable H\(_2\)O structure formed as a consequence of an auxiliary source of H\(_2\)O, but their approach and conclusions differ. The first group uses a DSMC method and finds that, assuming a secondary point source far from the arc, “they see no shock”; assuming a linear source centered on the shock and aligned on-axis, they obtain an arc. Rodionov et al. (1998), recomputing the first assumption by a NSE method, do find a “viscous” double-shock structure, resulting from the interaction between the supersonic flow from the main nucleus and that from the secondary source. The word “viscous” alludes to the fact that, because of the high gas rarefaction, the canonical double-shock (visible on Plate 13, left panels) is smoothed into a single structure. Figure 10 shows the re-
Fig. 10. Hyakutake’s coma on March 26, 1996, after Rodionov et al. (1998). (a) Computed H₂O number density (top); Mach number (center); computed OH number density (bottom); the observed OH arcs are similar to the observed C₂ arcs. (b) C₂ + dust (top); dust only (center); C₂ only (bottom). Notice the bright spot on the antisolar axis, visible in the center image.
sult of the computed interaction (assuming a secondary source representing some 10–20% of the total H₂O production). This solution is not unique, but it reveals the Hyakutake arcs as the first evidence of weak gas shocks in a coma. Even more importantly, it gives confidence in the validity of the previously described modeling results, acquired under much denser gas conditions: Near a dayside nucleus at 1 AU, the gas number density is \( \mathcal{O}(10^{13}) \text{ cm}^{-3} \), while in the stagnation region of the Hyakutake arcs it is computed to be \( \mathcal{O}(10^{8}) \text{ cm}^{-3} \) only (see Fig. 10). This corresponds to a Knudsen number \( \mathcal{O}(10^{-1}) \), quite convenient for the validity of the NSE approach (see Crifo et al., 2002a, 2003a). The quoted statement of Harris et al. (1997) that their DSMC result “does not see a shock” suggests some technical inadequacy in the implementation of their DSMC code.

7.2. P/Halley Near-Nucleus Dust Coma

The greatest of all possible challenges to a coma model was offered by the famous HMC synthetic image of Halley’s nucleus and its vicinity (Keller et al., 1995, and references therein). As stated earlier, this image has been unanimously — but without any supporting model computation — interpreted as “visual evidence” that “the activity is confined to localized regions.” However, “activity” means surface flux, and we do not “see” dust fluxes on images; we see dust column densities (more or less proportional to the local density, near to the surface). But near the surface, the dust is coupled to the gas. Observing dust in natural flows (e.g., snow blown on an icy surface, or desert dust blown by winds) is enlightening here: Dust accumulation often reveals flow stagnation regions, rather than high flux regions. There is no intuitive way to reliably infer flux from density; inferences must be checked by quantitative modeling.

The three-sources-on-a-sphere paradigm of Knollenberg et al. (1996) described in section 6.2 was offered in support of the classical interpretation, but we remind the reader of the weaknesses we have discovered: (1) the interaction between the outflows from the three regions cannot be neglected, and (2) the control of the outflow by the topography cannot be neglected — it is in all likelihood dominant. It is also evident that this model, which assumes a spherical surface, cannot be matched to the central part of the HMC images, which resolve a highly aspherical nucleus.

The “filament” map derived by Keller et al. (1995) using enhancement techniques is reproduced here in Fig. 11a. Knollenberg et al. (1996) proposed that each filament is the signature of a small-circular-inactive-area, in accordance with Plate 12. Here again, we must point out the weaknesses of this explanation: (1) The narrow collimation of the dust pencil obtained requires a strict cylindrical symmetry. This cannot be achieved by groups of closely positioned spots, as would be required; even the symmetry of an isolated spot will be ruined by the irregular topography of the surface where it lies. (2) Obviously, the Sun also cannot be placed simultaneously on the axis of many distinct spots. And, if one wanted to state that the gas outflow does not depend upon solar illumination, how then could one explain the vanishing activity observed on the nucleus nightside?

Crifo et al. (2002b) have offered a related, but self-consistent interpretation of the “filament” array. They observe that the visible-light HMC images are sensitive to all dust masses, and it is clear that the dust trajectories depend upon their mass; furthermore, nonspherical grains have trajectories totally different from spherical grains of the same mass. Thus, the simplest bet one should make when looking at

Fig. 11. Interpretation of the near-Halley-nucleus dust coma structures observed by the HMC camera (Crifo et al., 2002b). (a) Azimuthal brightness gradient map derived by Keller et al. (1994). (b) 0.91-µm-radius dust grains column density, computed by the present model assuming P/Halley’s nucleus to be strictly homogeneous with \( f = 0.6 \).
this map is that one sees the signatures of weak gas shocks, since this signature is due only to the difference in inertia between dust grains and gas molecules, irrespective of grain size and shape. This signature, being independent from dust grain size and shape, can be discovered by "numerical tracing": At the laboratory, one would inject calibrated dust (i.e., single-size, spherical grains) in the flow. Here, we can numerically simulate the injection of such calibrated dust. Furthermore, a single-fluid dust model can be used, even though it does not provide the exact dust density in the structures, since it does provide the correct structure position. Crifo et al. (2002b) consistently computed the overall gas outflow from the whole nucleus surface, using the nucleus shape derived from the images obtained by the Vega 1 probe by Szego et al. (1995), and assuming the nucleus to be (to start with) compositionally homogeneous.

Figure 11b shows the computed dust column density of spherical grains of 0.91-µm radius, computed with a single-fluid model. Detailed comparison with the HMC map (Fig. 11a) reveals the following: (1) The N1–N2 gradients are reproduced with an angular offset of about 20°. (2) All other gradients, in particular U2, U3, U4, N4–N8, and S2, are present in the model computation at the right position. This result is not significantly changed by variations in dust grain size. It is, on the other hand, very sensitive to the nucleus orientation — which is not known to better than a few degrees, hence the agreement seems surprisingly good.

So, for the first time in the history of cometary data analysis, dust coma structures are reproduced ab initio, i.e., without ad hoc assumptions. The physical meaning of this agreement is also clear: The coma structures seen are the signatures of gas flow discontinuities induced by the topography.

Comparisons of a similar kind (partially published in Szego et al., 2002) were made by the same authors with all Vega 2 images (the Vega 1 images carry little information on the near-nucleus coma) with even a more spectacular agreement than that obtained by Giotto — but here, one will notice that the resolution of the Vega 2 images is somewhat lower.

7.3. Gas Lightcurves

Gas lightcurves are obtained from summing the emission of molecules within a cylinder of radius R (in kilometers) (projected in the coma). One can compute that, if the gas velocity is on the order of 1 km/s, and the emission roughly isotropic, about one-third of the molecules were emitted over a time period R (seconds), and two-thirds over a time period of 1 d (case of OH at 1 AU from the Sun). Thus, for a real nucleus rotating with a period approximately equal to a few days, a 3D + t model of the gas coma coupled to a nucleus rotation model should be used to extract a correct Q(t) from the gas lightcurve. At this time, we know of no comet for which such a task was done.

The best prospect for seeing such a work performed is Comet Halley, since its complete nucleus shape is known and good lightcurve data are available.

The equations of rotational motion were integrated by Szego et al. (2001), based on the three nucleus orientations indicated by the imaging experiments during the flybys of Vega 1, Vega 2, and Giotto, with due allowance for the outgassing torque. Assuming a nucleus density 0.5 g/cm³, the torque is about a 5% large correction term for the equations of rotation. Due to this torque, the direction of the angular momentum vector is changed by −11° in longitude and 16° in latitude between August 31, 1985, and July 26, 1986, i.e., from 3 AU AP to ~2.8 AU PP. The computed rotation exhibits a basic rotation period ~2.2 d about the short axis of the nucleus, modulated by a 7.4-d long period. This model, despite its 7.4-d modulation, does not reproduce the basic ~7.4-d periodicity observed in ground-based observations (lightcurve, recurrence of jet patterns).

Another rotation model has been derived by Belton et al. (1991). It exhibits the requested rotation period of ~7.4 d (found to be about the long axis of the nucleus). This model, however, is obtained by flipping the orientation of the nucleus during the Vega 1 flyby with respect to the observed one. This seems to be due to the fact that the authors did not use the processed Vega 1 images, but only the raw images, much less constraining.

As already stated, Julian et al. (2000) and Szego et al. (2001) checked their P/Halley rotation models against a Q(t) derived from a Hase formula. One can therefore question the satisfaction the two groups express regarding their fits. This situation is not satisfactory, and work is in progress to find a solution for the rotation of the nucleus that satisfies all constraints (and hopefully all groups). It will then be possible to attempt an exact interpretation of P/Halley’s lightcurve.

7.4. Gas “Velocity Law”

In many cometary papers a so-called “Bobrovnikov’s velocity law” is advocated to evaluate “the coma gas velocity.” Keller (1954) quoted observations (without any technical details such as the names of the comets that were observed) of “jets and haloes” by Bobrovnikov that were later used by Whipple (1982) to derive the above-mentioned law. As far as we know, Bobrovnikov himself never published such a result, even decades after this report. Indeed, this “law” seems to mix up observations of a quite different nature: Anything moving is used. It is in particular quite unclear whether the observed motions are due to gas or to dust. Furthermore, many authors admit that this law (quite expectedly) is violated by their data. Last, but not least, one can read in Bobrovnikov’s (1931, p. 460) famous report on Comet Halley that “the assumption of a constant velocity of ejection is questionable. Data gathered in Chapter II indicate rather wide ranges of velocities near the nucleus, from several tenths to several km/sec.” The same statement is duplicated in the conclusion of the report (p. 581).

As evidenced by the model results presented above, there is nothing like a “gas velocity” in a coma; this velocity varies from point to point. This is observationally evident in the high-resolution observations of molecular radiowave emis-
sions (e.g., Henry et al., 2002). As for the heliocentric dependence, it is straightforward to show that, in the case of subsurface diffusion, the near-nucleus gas velocity is \( \propto r_h^{-1/4} \); in the case of surface sublimation, this dependence is negligibly small. For a discussion of outer coma gas velocities, see Combi et al. (2004).

7.5. Dust “Velocity Law”

In some favorable cases, e.g., the observation of a dust neckline structure (see Fulle, 2004), it is possible to derive the terminal dust velocity from an analysis of the image brightness. Most of the time, this is not possible. If one wants to derive a dust mass loss rate from an observed image brightness, it is then necessary to use assumed dust velocities. For this purpose, a “dust velocity law” is often advocated. Many variants of this “law” are circulated, including — as already pointed out in Wollis (1982) — some that are nothing but misprint propagations. Most variants postulate isotropic gas and dust production, and they all assume, quite unrealistically, that both the grains and nucleus are spherical. An expression summarizing the presently discussed gas dynamic simulations of (nonisotropic) spherical grain ejection from a spherical nucleus is given in Crifo et al. (1997a). Still, many heuristic parameters are present in this expression, for which different users will unavoidably choose differing numerical values when trying to represent real data. The resulting chaos in dust mass loss rate assessment was already pointed-out in Crifo (1991b) and will unavoidably persist until the missing information about dust grains is available.

More fundamentally, and as stated earlier, one must be careful to remember that dust is an ill-defined concept — or, if one prefers, the properties of real cometary grains are unknown. It is true that large-scale coma structures can be interpreted under the declared assumption of spherical grains, but that does not prove that the real grains are spherical. After decoupling from the gas, the grain motion is controlled by solar radiation pressure and solar gravity. The two effects do not distinguish the grain shape (a spinning irregular grain will exhibit an effective radiation pressure quite like a spherical one). Hence the success of outer coma fits gives no clue to what happens inside the gas-dust interaction region. Unfortunately, the ejection velocity of solids depends critically on their shape and even initial grain orientation at the surface (see Crifo and Rodionov, 1999), hence the ejection velocity of real cometary grains from a real cometary nucleus can simply not be predicted at present, save in a crude order of magnitude.

7.6. Dust Lightcurves

As opposed to the case of the gas, 3D + t outer coma dust models do exist, but to make their use rewarding, nuclei shapes are missing. Remembering that dust production rates are derived through the Afp method, it is easy to show that, while a spherical nucleus that would eject dust with perfect isotropy would build up a dust coma with a \( 1/t \) brightness dependence, the reciprocal statement is incorrect: Anisotropic dust-ejection patterns are capable of building up an apparently isotropic coma with brightness slopes widely dispersed around the value \(-1\), depending on the direction of observation (Fulle and Crifo, 1999). Hence, from this point of view, the production rates derived without knowledge of the nucleus shape should be treated with caution.

8. Physical Models of the Outer Coma Structures

We have seen that (1) steady-state gas and dust solutions cease to be valid at distances in excess of typically 1000 km and 100 km, respectively, and (2) the gas-dust interaction is completed before 100 km, so that the gas and the dust properties can then be computed independently; their outer coma structures need not bear any similarity. In both cases the following are required:

1. A physical (time-dependent) model must be used to compute the evolution of the gas and dust species distribution with distance to the nucleus. We will not review such models in the present chapter. (For the gas in a comet like Hale-Bopp, solving the NSE is one possibility.) The general observation is, however, of interest, that this evolution always involves structure smoothing, as substantiated below. Hence the more distant the observed regions, the less suitable they are, in principle, for inferring detailed nuclear properties.

2. Time-dependent boundary conditions must be specified at the distance where the steady-state solutions cease to be valid. We discuss in detail these boundary conditions, because, regardless of the quality of the outer coma model, the solution is just as valid as the boundary conditions used are acceptable.

8.1. Large-Scale Dust Structures

Plate 13 shows that the evolution of the structures from the vicinity of the nucleus to larger distances is very different for the gas and for the dust. The dust structures are quickly “frozen” into a fixed pattern (which changes with the rotation phase). The gas structures are not at all frozen. Large-scale dust structures can therefore be considered to be built up by free effusion from an “effective source” that is not the nucleus surface, but the terminal gas-dust acceleration surface. It is perhaps not evident that this effusion is a smoothing process, because it is a slow one: Smoothing is caused by the fact that, at each point of the distant coma, dust from all over the terminal surface arrives; only the unphysical myth of single-velocity, strictly radial dust ejection can lead to unique point-to-point dynamical connections between the coma and the surface of the nucleus. Even worse, dust from successive terminal surfaces during a sizable period (weeks) can possibly reach a given point. This evidently mixes up the terminal surface pattern.

As shown by Plate 13, this terminal surface has localized areas of enhanced dust flux, and indeed resembles Seka-
nina's (1981, 1991) “nucleus surface activity” maps. Hence, it is clear that what Sekanina and several other scientists using the same method determine is, in reality, not a nucleus surface distribution, but this effective source distribution. The difference in distances involved — O(100) km — is much smaller than the spatial resolution in ground-based coma images. One therefore easily conjectures why the authors are forced to turn “on” and “off” in an ad hoc way their “active regions”: This is most likely due to their attempt to fit a changing pattern (that at the terminal surface) with a supposedly fixed pattern (assumed to be the nucleus surface map).

For the nuclei of unknown shape, there is definitely no possibility of passing from such “terminal surface maps” to true “nucleus activity maps.” But for those with a known shape (at this time, only P/Halley), this is possible. Work in this direction is in progress within the frame of the International Space Science Institute “Halley” Team Work.

8.2. Large-Scale Gas Structures

For the gas, after the initial structuring by weak shocks, the persistence of the fluid behavior maintains a very efficient pressure smoothing (Plate 13). If the expansion was adiabatic, this smoothing would stop rapidly as $p \to 0$ when $T \to 0$. But the expansion is not at all adiabatic: The gas pressure continues to be significant because of heating by fast photodissociation products. Furthermore, the observed daughter species may have a distribution different from that of fast photodissociation products. Furthermore, the observed daughter species may have a distribution different from that of the precursors. This is all the more true of the conclusions regarding chemical inhomogeneity. The visionary prediction of Whipple (1950, 1951), that cometary activity is due to the solar-driven surface and subsurface sublimation of a solid nucleus dominantly made up of the ices of a few simple molecules and of dust, is still the basis of our efforts to understand comets today. For decades, the simple heuristic simplifications he made to derive orders of magnitude also remained the basis of most of the interpretations of the observational data. To some extent, this can be understood, as it was not possible to access detailed information about the nucleus and the dust. With the implementation of deep-space missions to comets, this situation is radically changing. It becomes possible to picture the cometary nuclei as real solar system objects, with their unavoidable extreme complexity. The price to pay is that, to understand them, the simple heuristic approaches must be abandoned in favor of the use of the full arsenal of the modern methods of applied physics. Even before the avalanche of data from cometary orbiters is at hand, enormous efforts should be dedicated to adapting these methods to the difficult problem of inferring properties of a real nucleus from observations of its surrounding gas and dust coma. We hope that the present chapter has convinced the reader that such an undertaking is (1) unavoidable, (2) challenging, and (3) potentially rewarding.

Acknowledgments. This work benefited from many successive yearly contracts from the French CNES and CNRS agencies, from partial support from the Austrian “Fonds zur Förderung der wissenschaft — lichen Forschung” under project P15470, and from the Hungarian OTKA grant T 32634. The International Space Science Institute in Bern is thanked for having supported an ISSI team meeting dedicated to the work on P/Halley discussed in this chapter.

REFERENCES

Benkhoff J. and Spohn T. (1991) Results of a coupled heat and mass transfer model applied to KOSI sublimation experiments. In Theoretical Modelling of Comet Simulation Experiments
31(4), 1128.


