Episodic aging and end states of comets

Zdenek Sekanina
Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena CA 91109, U.S.A.

ABSTRACT

It is known that comets are aging very rapidly on cosmic scales, because they rapidly shed mass. The processes involved are (i) normal activity—sublimation of ices and expulsion of dust from discrete emission sources on and/or below the surface of a comet’s nucleus, and (ii) nuclear fragmentation. Both modes are episodic in nature, the latter includes major steps in the comet’s life cycle. The role and history of dynamical techniques used are described and results on mass losses due to sublimation and dust expulsion are reviewed. Studies of split comets, Holmes-like exploding comets, and cataclysmically fragmenting comets show that masses of 10 to 100 million tons are involved in the fragmentation process. This and other information is used to investigate the nature of comets’ episodic aging. Based on recent advances in understanding the surface morphology of cometary nuclei by close-up imaging, a possible mechanism for large-scale fragmentation events is proposed and shown to be consistent with evidence available from observations. Strongly flattened pancake-like shapes appear to be required for comet fragments by conceptual constraints. Possible end states are briefly examined.

Keywords: comets, nucleus, fragmentation, aging, end state

1. INTRODUCTION

Comets are a category of celestial objects that show overwhelming evidence of spontaneous, almost reckless shedding of mass. This process of mass loss is observed as cometary activity, which generally consists of (i) sublimation of water ice and other, more volatile ices from the nucleus; (ii) dissociation and ionization of their gaseous products in the atmosphere and tail; and (iii) emission of dust (refractory material), which is entrained in the expanding flow of gases, accelerated away from the nucleus, and described by a size/mass distribution law and by thermal, scattering, and other physical properties.

Recent dramatic changes in our perception of the makeup of cometary nuclei, stimulated by in situ investigations of four periodic comets, include the relative amounts of the icy and refractory constituents in the nucleus and the ratio of their contributions to the total loss of mass. In the pre-Halley era, before 1986, the understanding of cometary nuclei had been strongly influenced by Whipple’s icy-conglomerate model depicting the volatile species as the main ingredient and contaminated by dust. Appropriately, Whipple’s model is often called a “dirty iceball” to emphasize the primary role of the ices. Their presumed prevalence in emissions from comets was supported by investigations as recently as the early 1980s. Delsemme,4 reinterpreting the results by others, derived for the dust-to-gas mass-production rate ratio in ejecta from two dust-rich comets a value near 0.7, while Greenberg,5 using his model for pristine comet composition, obtained for this ratio a value of 0.85. Comparing the dust-production rates calculated from his thermal infrared data with the water-production rates derived by others, Ney6 determined the dust-to-water mass-production rate ratio (which is always greater than the dust-to-gas ratio) for five comets observed between 1970 and 1976 to range from 0.10 to 0.83.

The data recorded by dust-particle detectors onboard the flyby spacecraft to Halley’s comet in 1986 implied a dust-to-gas mass-production rate ratio much higher than unity. The results of measurements conducted with a set of dust detection instruments on the European Space Agency’s mission Giotto to comet Halley, combined with data on the total output of gas, led McDonnell et al.7 to constrain the dust-to-gas mass ratio to a range from 1.3 to 3, while from their determination of the comet’s average atomic abundances of the elements Jessberger and Kissel8 found the ratio of 1.7.

These high dust-to-gas mass ratios for Halley’s comet are also understood in the context of one of the main imaging results, namely that only a relatively small fraction of the surface of the comet’s nucleus is active, the rest of it being covered by a dark, inert dust mantle (Ref. 9). Subsequent space missions and ground-based research have shown that, among comets, a similarly high dust-to-gas mass ratio and a mostly inert nuclear surface are apparently a rule rather than an exception.

Further author information: E-mail: Zdenek.Sekanina@jpl.nasa.gov, Telephone: 1 818 354 7589
A comet’s mass-loss rate of gas and dust peaks near, though usually not right at, perihelion, decreasing with increasing heliocentric distance. Models of cometary activity can be used to estimate the contributions from larger distances and to integrate the losses over the whole revolution about the Sun. For example, Sekanina\textsuperscript{10} estimated the total mass-loss rate of $\sim 2 \times 10^3$ g per revolution for comet 10P/Tempel, based on observations from its 1988 return to the Sun, when its activity peaked within days of the passage through perihelion. The nucleus of 10P/Tempel is fairly large for a short period comet and was approximated by a triaxial ellipsoid,\textsuperscript{10} whose dimensions were 16.4 km by 9.8 km by 7.0 km. At an assumed bulk density of $\sim 0.4$ g/cm$^3$, its mass is estimated at $\sim 2 \times 10^7$ g, so that the comet lost only 0.01 percent of its mass per revolution ($\sim 5.3$ years) and could, at this rate, still survive some 10,000 revolutions or more than 50,000 years. With a few notable exceptions, such as Halley’s comet, most periodic comets have nuclei less massive but they are also less active. Based on Lisse’s\textsuperscript{11} results, a more typical periodic comet, like 9P/Tempel, loses only $\sim 10^2$ g per revolution.

For periodic comets, both the mass-loss rate and the dust-to-gas mass ratio can vary significantly from return to return. Schleicher\textsuperscript{12} showed that between 1983 and 2005, during a span of four revolutions, comet 9P/Tempel’s production of water (based on his hydroxyl data) decreased by a factor of 2.4, while the dust production was down by only a factor of 1.3; the dust-to-gas mass ratio thus increased by a factor of almost 2. This comet differs from 10P/Tempel in that both its gas production and dust production peak 1–2 months before perihelion and even the 1983 production rates were below 10P/Tempel’s 1988 values.

2. COMETARY AGING STUDIED DYNAMICALLY

If comets age by sublimating ice and emitting dust, it would seem that in order to estimate for how long a comet can last, all that is needed is to sum up these two contributions. Unfortunately, it turns out that just as any other celestial body, and in fact much more so, comets are exposed to cosmic processes that can shorten their life span significantly, sometime cataclysmically.

Before Whipple’s model was developed in the early 1950s, the mass loss rates of comets were thought to be much lower than nowadays. Traditional chemical species identified spectroscopically in the optical spectrum, such as C$_2$, CN, C$_3$, etc., are known today to represent only a very small fraction of the total mass of gas released by comets. Similarly, microscopic particles detected since long ago by their optical scattering of sunlight were believed to carry most of the mass of nonvolatile material, a notion that unfavorably affected the understanding of cometary dust until the 1980s (Sec. 1). It was Whipple who predicted, primarily on the basis of his dynamical arguments, that the mass loss rate from comets must be orders of magnitude higher than indicated by contemporaneous observations, and this bold idea was not confirmed until vast hydrogen clouds about comets were observed for the first time with the satellite-based imagers some 20 years later. Dramatic broadening of the spectrum of wavelengths over which comets began to be observed more recently only further strengthened the validity of this remarkable prediction.

2.1 Outgassing-Driven Nongravitational Perturbations of Motions of Periodic Comets

It has been known since the 1820s that the orbital motions of comets, unlike those of asteroids, may fail to strictly obey the law of gravitation. The deviations in the orbital period are very small, on the order of 0.01 percent or less, but readily detectable in the orbital period. Encke\textsuperscript{13,14} and his followers tried for nearly a century to prove that his hypothesis of interplanetary resisting medium was compatible with the slightly accelerated motion of a comet that carries Encke’s name and orbits the Sun once in 3.3 years. Until the time of his death in 1865, Encke\textsuperscript{15} did not have the slightest doubts about the existence of the resisting medium, yet efforts to continue his orbital investigations were for insurmountable difficulties intrinsic to the hypothesis eventually abandoned several decades later.\textsuperscript{16} Bessel’s\textsuperscript{17,18} brilliant idea of recoil action from ejecta of Halley’s comet on its center of gravity as a trigger of these nongravitational perturbations — a paradigm that is in line with the modern understanding of comets — had not been seriously entertained by comet scientists for more than a century, until independently proposed by Whipple\textsuperscript{1} in his icy-conglomerate model.

In one detail, Bessel’s and Whipple’s solutions differ. Bessel attributes the nongravitational effect (acceleration versus deceleration) to a perihelion asymmetry in the mass loss, the comet being accelerated if the preperihelion production is higher and vice versa. On the other hand, Whipple attributes the nongravitational effect to sublimation lags of a spinning nucleus, the comet being accelerated when it rotates in a sense opposite its orbital motion and vice versa. Thus, Bessel explains the effect regardless of the sense of rotation and/or sublimation lags, while Whipple explains the effect when the comet’s production is symmetric relative to perihelion.
Systematic investigations of the effects of outgassing-driven nongravitational forces on the orbital motions of periodic comets were initiated in the late 1960s by Marsden\textsuperscript{19,20} and continued by him and his collaborators through the 1970s and beyond. One of the highlights of this work was the introduction, in 1973, of a Style II system of nongravitational parameters\textsuperscript{21} that is still in use worldwide today as a standard computational technique for high-precision orbit-determination efforts. The Style II set of parameters is determined from the nongravitational terms incorporated in the equations of motion as an additional orbital perturbation, using a physically sound sublimation law for water ice, which is symmetrical with respect to perihelion, and an RTN right-handed coordinate system (radial, transverse, normal) aligned with the comet’s orbital plane and the Sun-comet direction as the reference orientations for the components of the nongravitational acceleration. The effect of the nongravitational force is in general described by three parameters that are derived from a standard-type least-squares iterative differential-correction orbital solution together with the usual set of elements: the parameters describe the magnitudes of the nongravitational acceleration normalized to 1 AU from the Sun in the direction away from the Sun, $A_1$, in the direction perpendicular to the radial direction in the orbital plane ahead of the comet, $A_2$, and in the direction normal to the orbital plane, $A_3$. The customarily used unit for these parameters is $10^{-8}$ AU/day$^2$. The Sun’s gravitational acceleration at 1 AU, 0.593 cm/s$^2$, is equal to $2.959 \times 10^4$ in these units.

Experience with applying this new computational technique soon showed that two apparitions of a comet could always be linked by a purely gravitational solution. In order to derive the nongravitational parameters, a minimum of three apparitions was needed. The transverse component’s parameter $A_2$ was, as a rule, determined most accurately, while the normal component of the nongravitational acceleration was almost always negligibly small and indeterminate, so that one could assume $A_3 = 0$. The parameter $A_1$ of the radial component could often be derived with only low accuracy, which at first sight seemed puzzling. The nongravitational effect in the orbital period was determined by the transverse component: the motion of Encke’s comet has been accelerated because its $A_2$ has always been negative. The enormous difficulty that Encke’s followers experienced were due entirely to a systematic decrease in the absolute value of the transverse component of the nongravitational acceleration that the comet was subjected to during its observed history. Until the return of 1855, near the end of Encke’s era, the parameter $A_2$ was confined to a relatively narrow range of values,\textsuperscript{22} from $-0.035$ to $-0.040 \times 10^{-8}$ AU/day$^2$, which explains why Encke felt so confident about his hypothesis of resisting medium. By the time Backlund\textsuperscript{16} completed his orbital computations extending the interval to 1908, the transverse acceleration dropped\textsuperscript{22} to less than one half of its value before 1855, forcing Backlund to give up. Currently, the comet’s parameter $A_2$, though still negative, is some 30 times smaller than during Encke’s times.

There could be three causes for this systematic trend: (i) the comet’s production curve may have become more symmetrical relative to perihelion (consistent with Bessel’s version); or (ii) the comet’s rotation axis may have gradually tilted closer to the orbital plane or the sublimation lag grew smaller (consistent with Whipple’s version); or (iii) the momentum transferred to the nucleus by the outflowing matter decreased with time on account of diminishing activity of the comet, which would be consistent with both Bessel’s and Whipple’s models. Point (iii) certainly could not explain the whole effect, since the comet’s activity has not declined dramatically over the past two centuries. In fact, Kamél\textsuperscript{23} stated that the activity remained steady for at least 150 years. He found that the light curve of Encke’s comet, which is known to be intrinsically brighter before perihelion, has more recently shifted, reaching its peak nearly at perihelion. On the other hand, Whipple and Sekanina\textsuperscript{24} were able to fit the variations in the comet’s nongravitational acceleration from 1786 to 1977 using their precession model. It is therefore possible that one observes a combined effect of causes (i) and (ii).

These are issues that are critical for our understanding of the evolution of both the surface and interior of nuclei of comets, especially periodic comets,\textsuperscript{25} as they relate to the processes of activation, deactivation, dormancy, reactivation, and extinction of individual discrete emission sources that are known to supply much, if not all, of cometary activity. Bessel’s\textsuperscript{17,18} interpretation received a boost from investigations of a correlation between the perihelion asymmetry of the gas production curve and the nongravitational effect in the orbital period among dozens of periodic comets. These studies showed that the short-period comets of larger perihelion distance have a tendency to be less dust covered and/or to have smaller nuclear dimensions and that those of them recently perturbed into orbits of a much smaller perihelion distance were subjected to the strongest nongravitational effects. Comparison with comets that resided in orbits with a small perihelion distance over extended periods of time and were subjected to weak nongravitational perturbations implies a possible evolutionary track for the formation and growth of inert dust mantles on low-activity periodic comets.

Much insight into the physical meaning of the nongravitational parameters has been gained from the dynamical analysis of the Style II nongravitational law.\textsuperscript{29} Although this law, by virtue of its symmetry with respect to perihelion, obviously fails to match temporal variations in the momentum transferred to the nucleus by a gas flow that is asymmetrical relative
to perihelion, this law accounts for the nongravitational perturbations of the orbital elements upon the integration over the entire orbit about the Sun. Even though, for example, the radial component of the transferred momentum contributes to the nongravitational perturbation of the orbital period, the latter depends only on \( A_2 \) and not \( A_1 \) because of the law’s perihelion symmetry. These considerations further indicate that the parameter \( A_1 \) actually measures the nongravitational effect on the comet’s line of apsides: a positive \( A_1 \) implies its regression, a negative \( A_1 \) its advance. The general rule regarding the sign of the parameter \( A_2 \) that follows from the baseline model, which applies to the case of a single emission source on the rotating nucleus with no sublimation lag, can be formulated as follows: \(^{29}\) The parameter \( A_2 \) is positive and the gas production curve peaks after perihelion when the emission source is on the hemisphere of the rotation pole that is sunlit and on the trailing side of the nucleus at perihelion or when the source is very near the equatorial plane or on the other hemisphere than the pole that is sunlit and on the leading side at perihelion, and vice versa. In this baseline model the nongravitational parameters are independent of the sense of rotation. The model can be readily generalized to include sublimation lags (which invalidate the above rule) and to any number of emission sources.

Of particular interest for studying sudden activity developments on the nucleus are periodic comets that have undergone the so-called discontinuous orbital anomalies. These comets exhibit very dramatic changes in the nongravitational parameters, such that it is impossible to link a few consecutive apparitions in a single orbital run. These discontinuities offer probably direct evidence of a new emission source’s sudden activation or an existing active source’s sudden deactivation or extinction. Since any emission source can be envisaged as consisting of a number of point-like sources, it is in principle possible to examine the distribution of emission sources on the nucleus once its shape and rotation vector are known. Such a technique has actually been developed\(^ {29}\) and applied to several periodic comets with discontinuous orbital anomalies.

It is apparent that each emission region, whether it releases gas and dust from as little as a fraction of 1 percent or as much as \(~\sim\)10 percent of the nuclear area, and whether it is located on the surface itself or occupies a cavity below the surface from which the fl w of emitted mass is seeping to the surface through an orifice the life span of an emission source amounts to a certain phase in the comet’s evolution. The source’s extinction represents a minor event of episodic aging; thus, the so-called normal activity of a comet is not a truly continuous process as it may appear on a large scale.

### 2.2 Nongravitational Perturbations of Single-Apparition Nearly-Parabolic Comets

During the 1970s, it also became clear that besides the periodic comets, the nongravitational orbit-determination technique would fin applications also among single-apparition comets moving in nearly-parabolic orbits, although their number would be low. Yet, this was a surprise, considering that a minimum of three apparitions (two full revolutions about the Sun, typically a decade or longer) was necessary in the case of periodic comets. The eight originally determined nongravitational orbits\(^ {21,30}\) still make up some 40 percent of the whole sample of 21 comets before the year 2000.

Because of no apparition linking involved, the parameters \( A_1 \) and \( A_2 \) show, respectively, the true contributions from the radial and transverse components. The parameter \( A_1 \) came out always well determined, positive as expected, and, with one exception, larger, often by a factor of 10 or more in absolute value, than \( A_2 \), which is positive in nine and negative in 12 cases. There is a possible relationship with fragments of the split comets (Sec. 3.1), as seen from the histograms in Fig. 1. Although the peak of the \( A_1 \) set appears falls between \(+1\) and \(+2\) units of \(10^{-8}\) AU/day\(^2\); all single-apparition comets whose motions were fitte by a purely gravitational solution would make a huge peak at \( A_1 = 0 \). The three comets with \( A_1 < +1 \times 10^{-8} \) AU/day\(^2\) are probably near the detection threshold of the nongravitational perturbations.

### 3. Fragmentation

The true rate of mass loss from comets is higher than shown in Sec. 1, sometimes dramatically so, because of another process to which they are subjected — sudden fragmentation of their nuclei. Since only recently has it been recognized that comet fragmentation is an omnipresent process, which can occur at any heliocentric distance, which dramatically affects the properties and behavior of the comet’s atmosphere by saturating it temporarily with dust, and which may significantl accelerate the process of aging.

Although a fragmentation event as such can only be observed with in situ instruments onboard a rendezvous space mission hovering, at the critical time, in the proximity of the fragmenting nucleus, its signatures are detected by Earth-based observers fir as an outburst and days or weeks later as two or more discrete nuclei or components — one principal and one or more secondaries, boulder-sized or larger active fragments or their clusters — that replace the single nucleus and recede steadily from each other. The time of a fragmentation event, which can be computed by modeling the recession
Figure 1. Distribution of the radial nongravitational parameter $A_1$ of 21 single-apparition comets (lower part), which reached their perihelia before the year 2000 and which required the introduction of the nongravitational terms in the equations of motion to successfully fit their orbits, are compared with the distribution of the differential decelerations $\gamma$ for 37 secondary (companion) nuclei of comets observed double or multiple before 2000 (upper part, inverted). $N$ is the number of cases. A unit of $10^{-8}$ AU/day$^2$, customarily used for the parameter $A_1$, corresponds to 3.38 units of $10^{-8}$ the Sun’s gravitational acceleration, commonly used for the decelerations $\gamma$. Distinct similarity between the two distributions is noted. The entry at $A_1 = +9.26$ units of $10^{-8}$ AU/day$^2$ is comet C/1999 S4 (LINEAR).

It is only under extremely favorable circumstances that the physical evolution of a fragmented comet can be examined from Earth in greater detail. Comet 73P/Schwassmann-Wachmann, a Jupiter family comet with an orbital period of 5.4 years, offered us recently an opportunity for close-up examination of the properties and behavior of its fragments. The comet was seen to be multiple in both its 1995/1996 and 2000/2001 returns to the Sun, and it was recognized as a candidate for a potentially spectacular show of fragments during its very favorable return in 2006, when it approached Earth to only 12 million km. The wildest expectations were fulfilled as more than 600 fragments, looking like minicomets (Fig. 2a), were observed well enough to be assigned official designations, with an unofficial fragment count in the hundreds. Some fragments were particularly active (Fig. 2b), and their progressive fragmentation was accompanied by recurring outbursts (Fig. 3) apparent on the light curves. All fragments separating directly from the original comet’s nucleus made up a population of fragments of the first generation. All fragments separating directly (though not necessarily simultaneously) from these fragments made up fragments of the second generation, etc. A sequence of events giving birth to two or more generations of fragments is said to form the process of cascading fragmentation.
Figure 2. Comet 73P/Schwassmann-Wachmann as a product of cascading fragmentation. Left: (a) The train of nuclear fragments, imaged on May 4–6, 2006 by the Spitzer Space Telescope in the thermal infrared (24 \mu m) and spanning a length of \( \sim 10 \) million km (37° in projection onto the plane of the sky). The brightest, principal fragment C is located near the leading end, toward the upper right. The second brightest fragment, B, and other companions align along the projected orbit behind C, to the lower left of it. (Image credit: Spitzer Space Telescope.) Right: (b) Fragment B with a cluster of nearby, recently released subfragments, located nearly along the tail, to the lower right of their parent fragment are seen in this high-resolution image taken by Carl Hergenrother with the 180-cm f/9 Vatican reflector on Mount Graham in Arizona on May 1, 2006, using a Harris R filter. The diagonal of the field of view is about 35,000 km at the comet. More than 60 subfragments were detected in the original image. (Image courtesy of C. W. Hergenrother.)

Because events of fragmentation (or splitting) involve large objects (Sec. 3.1), they represent, unlike the events of activation and deactivation of discrete outgassing sources, major steps in the episodic loss of mass. The rate of aging over time intervals much longer than the orbital period depends on the rate of splitting, which unfortunately is poorly known, because the database is rather incomplete. Chen and Jewitt\textsuperscript{34} estimate an average rate of splitting at about 0.01 per year per comet, which probably is a crude lower limit.

Figure 3. Detail of the light curve of fragment B of comet 73P/Schwassmann-Wachmann from mid-March to mid-June 2006, exhibiting four outbursts accompanying discrete fragmentation events. Following the peak of each outburst, the brightness, expressed in terms of a visual magnitude \( H_{\Delta} \) normalized to a unit geocentric distance, always subsided more slowly than it had risen before the peak.
3.1 Split Comets and Differential Deceleration of the Motion of Their Companion Nuclei

Because of the extremely weak gravity field of comets, splitting of the nucleus leads typically to fragments that begin to travel through space independently of one another with very low separation velocities, on the order of 1 m/s. Observations of close double and multiple comets show invariably that, shortly after an event of nontidal splitting, the leading fragment is the principal (usually the brightest) nucleus, with the secondary, fainter fragments (called hereafter companions) trailing the main body in its tail, appearing sometimes like beads on a thread of worsted.

This peculiar arrangement suggests that the companions move, just like dust particles in the tail, in a gravitational field that is slightly weaker than the Sun’s field and that therefore their motions are affected, relative to the motion of the principal nucleus, by a differential deceleration. The force that diminishes the strength of the gravitational field for dust-tail particles is solar radiation pressure. For companion nuclei, which for a limited period of time are sources of activity, the radiation pressure will not do, but an outgassing-driven non-gravitational force, the same one that plays a major role in Whipple’s icy-conglomerate model, will. The smaller the companion fragment, the greater the antisolar force per unit mass — and therefore the differential deceleration, because the momentum transferred from the outgassing varies as a square of the companion’s characteristic dimension, whereas its mass varies as a cube. Thus, in the string of beads of a fragmented comet the most massive companions are nearest the principal nucleus, the least massive are farthest from it.

As the time from breakup increases, the separation distance between the principal nucleus and companions grows and so does the deviation of the latter from the antisolar direction of the principal nucleus, in correspondence with the conservation of orbital angular momentum law. Eventually, long after the breakup, the companions align along the comet’s projected orbit behind the principal nucleus. This evolution is illustrated in Fig. 2, where the fragments separated from one another long before the Spitzer Space Telescope image (left panel) was taken align as described, at an angle of about 30° with the tails. By contrast, a close-up image in the right panel shows that the subfragments released from fragment B (the bright spot to the upper left) shortly before the image was taken, during the early April outburst (Fig. 3), cluster close to the antisunward direction, near the tail (lower right of the parent). The projected orbit is to the upper right of the parent.

Most objects in the left panel of Fig. 2, including companion B, are first-generation fragments of the original parent comet. By contrast, the objects that clump in the right panel are second- and higher-generation fragments. It is thus observed that the process of fragmentation keeps repeating itself quite rapidly with ever smaller pieces, until the disintegration of some fraction of the original comet is completed. This is the cascading nature of the process. From this it follows that the life span of fragments statistically varies with their size. Most second- and higher-generation fragments in the right panel of Fig. 2 do not survive for more than a few days or weeks. The principal fragment almost always survives.

The separation parameters, including the differential deceleration $\gamma$ assumed to vary as an inverse square of heliocentric distance, have been derived for numerous split comets from a fragmentation model. A set of $\gamma$ values for 37 companions of 28 comets that split before the year 2000 is plotted in a histogram in the upper part of Fig. 1. Excluded from the plot are the sungrazing comets, for which $\gamma$ is inherently correlated with the separation velocity and cannot be determined. The distribution of $\gamma$ shows a remarkably strong correlation with the distribution of the nongravitational parameter $A_1$ for the single-apparition nearly-parabolic comets that require the introduction of nongravitational terms into the equations of motion, suggesting that the single-apparition comets with a clearly-positive $A_1$ are companion nuclei of split comets whose principal fragments are unknown, possibly because they arrived many centuries or perhaps even millennia ago. However, it should be noted that some of the comets with the clearly-positive $A_1$ were bright, very active objects and that some may have arrived from the Oort cloud, tens of thousands of AU away, both points complicating the issue.

A property of companion fragments of the split comets that is critical for estimating their sizes and masses is their shape. The derived decelerations $\gamma$ are in fact so high that if the companions were nearly spherical, they could not be as bright and active as indicated by observations. On the other hand, their light curves imply much lower decelerations than found from their relative motions. This contradiction could be removed only by requiring that the fragments be pancake-shaped, as if they were sliced off from their parent nucleus. In this scenario, each fragment’s brightness and activity are determined primarily by its large cross-sectional or base area, whereas the deceleration is determined primarily by its thickness, the minimum dimension. For a fragment separated from its parent comet 2–3 km in diameter and receding from it with a differential deceleration near the peak in Fig. 1, some 5–8 units of $10^{-5}$ the Sun’s gravitational acceleration, this model implies an average thickness of 0.1–0.2 km, a base area of 1.5–2 km$^2$, and an edge-on cross-sectional area of $\sim$0.1–0.3 km$^2$. With an assumed bulk density of $\sim$0.4 g/cm$^3$, this fragment’s mass would be between 0.6 and $1.6 \times 10^{14}$ g or $\sim$100 million tons. Fragments with higher decelerations would have correspondingly smaller sizes and masses.
3.2 Holmes-Like Exploding Comets with Rapidly Expanding, Sharply-Bounded Dust Halos

Comet 17P/Holmes is a faint Jupiter family comet orbiting the Sun at present between 2.05 and 5.2 AU with a period of about 7 years. Until very recently, the only unusual circumstances about it were its discovery, in 1892, during an outburst and another outburst some seven weeks later. The comet was visible by the naked eye for a limited period of time on both occasions, its brightness increasing by about a factor of 100 soon after the beginning of the second event. After the discovery apparition, the comet was so feebble that it was eventually lost for seven returns to perihelion, between 1906 and 1964. Skepticism prevailed about the comet’s ever again mimicking its 1892–1893 behavior in any significant measure. Yet, on October 23, 2007, after 115 years, the comet exploded again, this time brightening by a factor of about 400,000! In terms of both the amplitude and the peak intrinsic brightness, this was the most powerful event of this kind ever observed. It was most fortunate that only a few hours after it began, the developing megaburst was detected by J. A. Henriquez Santana, Observatorio Atlante, Tenerife, and shortly afterwards it was picked up by several alerted observers. Nearly 10 hours after the onset, Henriquez reported the comet brightening by a factor of 1.6 per hour. At this rate, the peak brightness would have been reached after 28 hours, but because the rate was gradually decreasing before a peak was reached, the duration of the active phase is estimated to have lasted longer by some 15 or so hours.

Observationally, a megaburst begins, like other, less powerful outbursts, with the appearance of a stellar nucleus and its steady expansion into a disk of nontrivial dimensions. The bright nucleus is the first manifestation of the expanding cloud of ejecta, mostly microscopic dust particles. The active phase, during which the amount of dust in the atmosphere keeps increasing, can last from a fraction of a day to many days. The steady expansion of the dust cloud or halo provides an opportunity to determine the event’s onset time. As a bonus, the rate at which the cloud grows in size measures directly the ejecta’s expansion velocity, which is diagnostic of the physical processes involved (Sec. 4).

The evolution of the products of the 2007 megaburst of comet 17P/Holmes could readily be followed for many months after the event’s active phase had terminated. The halo’s expansion, illustrated in Fig. 4 by a sequence of images taken by P. Vasey two to four weeks after the onset of the megaburst, also continued for months. As the integrated effect of solar radiation pressure on microscopic dust particles kept increasing, the halo’s boundary on the antisunward side became gradually more diffuse with time. The rest of the boundary, however, continued to be remarkably sharp, catenary-like, and — except in its central region — essentially featureless. The nuclear condensation had a tendency to grow in size, but at a much slower rate than the halo, and to display a steadily lengthening tail in the antisunward direction, another effect of solar radiation pressure.

After the extremely steep brightness increase during the megaburst’s short, active phase, the comet’s light curve reached a plateau almost two days after the onset and then began a very slow, gradual decline over a period of nearly six months. The comet was widely seen as a naked eye object for at least two months, while experienced observers could see it without using a telescope or binoculars for yet another month. Although the observed integrated brightness was subsiding due primarily to the comet’s increasing heliocentric and geocentric distances, the surface brightness was dropping more steeply on account of the expansion, so that the comet appeared to fade more rapidly than it actually did.

Analysis of the data on the megaburst of comet 17P/Holmes available by mid-January 2008 indicated that the halo’s expansion velocity was 0.50 km/s, thus showing that a significant fraction of the halo’s cross-sectional area, which exceeded 57 million km², was supplied by microscopic dust particles — micron- and submicron-sized grains. Use of an appropriate mass-distribution law led to a conservative estimate of $10^{14}$ g for the total mass of dust in the halo. This is an enormous amount, more than 1 percent of the comet’s estimated mass and unrivaled by outbursts experienced by other comets. Even the two events that 17P/Holmes underwent in 1892–1893 can only marginally compare with this megaburst, as the mass of dust injected into the atmosphere during both of them is estimated at $0.3 \times 10^{14}$ g. Only very recently has it been established that Halley’s comet experienced the same kind of event in January 1836, when it flare up dramatically at the time of initiation of a similarly bright, rapidly expanding and sharply-bounded halo. It is estimated that a total of $0.6 \times 10^{14}$ g of dust was injected into Halley’s atmosphere during that episode.

The nearly featureless outer regions of the halo in Fig. 4 support the notion that the megaburst was an event of global proportions on the scale of the comet’s nucleus, because the absence of morphology suggests that the material injected into the atmosphere originated from a source (or sources) of great extent. It is proposed that these giant outbursts, with the mass of injected dust from $10^{13}$ to $10^{14}$ g be called Holmes-like explosions and the comets observed to experience them Holmes-like exploding comets. Unlike 17P/Holmes, Halley’s comet is not a Jupiter family comet. Yet, significantly, all four events referred to above occurred along the post-perihelion leg of the orbit.
3.3 Cataclysmically Fragmenting Comets

From Secs. 3.1 and 3.2 it follows that regardless of the fact that a sizable fraction of mass during an event is lost, whether via cascading fragmentation or a Holmes-like explosion, the main object (principal fragment, exploding comet) survives and continues to orbit the Sun. There is a category of comets that fade and eventually vanish before the eyes of the observers on a time scale of several days to weeks. These rapid physical changes result in the tail becoming the object’s brightest part, surviving the head. The fading is usually accompanied by the disappearance, in a matter of days, of the nuclear condensation. At the same time, the coma often expands and becomes progressively elongated and its surface brightness decreases until the comet’s whole head is gone.

It was remarked more than 20 years ago that the terminal changes experienced by the comets of this category bear a resemblance to the physical behavior of split comets’ companions, and that the two phenomena are likely to be governed by the same or very similar processes. This conclusion was reached at a time when no high-resolution observations were available for either category of objects and when the known numbers of such objects were low compared to the current numbers. In those days, the quest for better understanding the vanishing comets was thwarted by inherent difficulties in observing them, starting with the dearth of reliable ephemerides and ending with the unpredictability of physical change on short time scales and sudden demise. It was nearly impossible to get immediate access to more sophisticated observational techniques of the time and, in the absence of the internet, to organize a coordinated campaign to monitor suspect objects.

A breakthrough was achieved in 2000. Comet C/1999 S4 (LINEAR), discovered 10 months before perihelion at a heliocentric distance of 4.3 AU began to show clear signs of anomalous behavior in early July 2000, three weeks before perihelion. An improved orbit determination based on astrometric observations up to July 6 indicated that the comet’s motion was being influenced by outgassing-driven nongravitational forces. At about the same time, observations made with a STIS detector of the Hubble Space Telescope (HST) on July 5–7 showed that the comet was in outburst on the first date, while at least one fragment was detected at a projected distance of ~460 km from the nucleus on the last date. From the parallel spectroscopic observations, the hydroxyl production rate on July 6 indicated that the comet’s water-outgassing area was at least 5 km², while carbon monoxide was found to be strongly depleted relative to water in comparison with abundances for other comets. Unexpectedly, D. G. Schleicher and C. Eberhardt’s hydroxyl abundance measurement on July 13, one week later, required a water-outgassing area of only 1 km².
By mid-July the comet’s unconventional behavior was already amply documented. Nightly imaging monitoring with a 100-cm Jacobus Kapteyn Telescope in Canary Islands began on July 23. The brightest part was highly condensed and teardrop-shaped the first night. Its morphology was unchanged but fainter by more than 1 magnitude the next night. The condensation became rhomboidal in shape, with its brightest part looking like a short, fat cigar the third night. Subsequently, the condensation continued to fade and its elongation to grow at a projected rate of ~40 m/s, with a very flat brightness distribution. By July 27 the condensation was no longer visible. These developments were generally confirmed by images taken at other observatories. The cigar was made up of expanding dust and its increasing length was diagnostic of the particles’ highest radiation pressure accelerations. It eventually developed into a narrow synchronous tail, whose origin was traced to the outburst on July 22\textsuperscript{8} ± 0.2 UT and whose mass was estimated at 4 \times 10\textsuperscript{11} g. Schleicher and L. Woodney’s\textsuperscript{45} hydroxyl data obtained on July 29 showed that the water-outgassing area decreased to a mere 0.2 km\textsuperscript{2}.

No individual fragments were observed in the cigar in any images taken with large ground-based telescopes. It looked as though the comet “blew apart” completely. But then a request for observing time on the HST was granted and breathtaking images were obtained on August 5: the tip of the tail — the location of the missing head — was populated by a cluster of more than a dozen minicomets. These were truly “orphan” fragments, none substantially brighter than any of its peers. The next night, a search for orphans became a task for the ANTU, one of the four units of the Very Large Telescope (VLT) at the European Southern Observatory in Chile, and the ANTU became the only ground-based telescope that succeeded. The mystery of vanishing comets has been solved, and the culprit was again — fragmentation.

The difference between images taken a few nights apart with a 250-cm and an 820-cm telescopes is illustrated in Fig. 5. Besides the detection, the HST and VLT images made it possible to determine that some orphan fragments separated from one another in early to mid-July and that their differential decelerations were on the order of 25–50 units of the Sun’s gravitational acceleration, comparable with the comet’s radial nongravitational perturbation determined by the orbit-determination technique (Fig. 1), which was 31 units of the Sun’s gravitational acceleration. Different orphan fragments were born at different times, but the rate of fragmentation may have accelerated in late July. Ten fragment condensations in the HST images were analyzed to reveal central, point-like objects, and the results showed that with an adopted geometric albedo of 0.04 and a phase slope of 0.04 mag/deg, the brightness peaks corresponded to the cross-sectional areas for these orphan fragments between 0.002 km\textsuperscript{2} and 0.011 km\textsuperscript{2}, their total amounting to ~0.07 km\textsuperscript{2}.

![Figure 5. Cataclysmic fragmentation of comet C/1999 S4 (LINEAR). Left: (a) Processed image of the headless tail taken by R. Corradi and N. O’Mahoney, Isaac Newton Group of Telescopes, with the Wide Field Camera of the 250-cm Isaac Newton Telescope at the Roque de Los Muchachos Observatory, La Palma, Canary Islands, Spain, on August 1, 2000. The diagonal of the field of view covered is 140,000 km at the comet. The Sun is to the right and slightly up. The comet’s missing head is at the tip of the tail to the lower right. With one exception, this was all that could be detected with any ground-based telescope. (Image credit: M. Kidger, Instituto de Astrofísica de Canarias.) Right: (b) Processed composite of three images taken with the ANTU, one of the European Southern Observatory’s four 820-cm Very Large Telescope (VLT) units at the Paranal Observatory, Chile, on August 6, 2000. The diagonal of the field of view covered is 47,000 km at the comet. The Sun is to the upper right. The square and the diamond are predicted ephemeris positions of the original nucleus from two orbital solutions available at the time. The VLT was the only ground-based telescope that resolved the comet’s head into the orphan fragments, a total of 16 in this image; the letters A–R mark their positions. Several days later, no trace of the fragments was detected with the VLT. (Image credit: European Southern Observatory.)](image-url)
It is concluded that the disappearing comets are cataclysmically fragmenting comets and that — as is clear from comparison of Fig. 2b with Fig. 5b — there is no fundamental difference between a cluster of subfragments (higher-generation fragments) released from a companion of a split comet on the one hand and a cluster of orphan fragments from a cataclysmically fragmenting comet on the other hand, except that the subfragments accompany the surviving parent body, while the orphan fragments are on their own. Both clusters are products of cascading fragmentation and, seen under low resolution, should look like an expanding coma whose elongation increases and surface brightness decreases with time. Eventually, this coma must disappear, as observed. The contradiction between the large water-outgassing area of C/1999 S4 and the anomalously high outgassing-driven nongravitational deceleration is solved by the comet’s very uneven, pancake-shaped dimensions, similarly to the case of companion nuclei (Sec. 3.1). From the summed-up cross-sectional area of 10 fragments of C/1999 S4, the total area of all orphan fragments is estimated at 0.1 km². Since the cross-section diminishes with time (see the caption to Fig. 5b), the pancake’s original minimum cross-sectional area must have been greater than 0.1 km², which compares favorably with the minimum cross-sectional area of companion fragments, estimated at up to 0.3 km² in Sec. 3.1. The companions’ maximum cross-sectional area, estimated at a few km² in that section, is nearly identical with the water-outgassing area of C/1999 S4, apparently this comet’s maximum area in an early stage of the disintegration process. In other words, C/1999 S4 — and, by extension, at least some other cataclysmically fragmenting comets — could be companions to the principal fragments of comets that had split in the distant past. To make this inference consistent with the finding that C/1999 S4 may have its (ill-defined) aphelion near the Oort cloud, the involved time span may be extraordinarily long. In any case, as in Sec. 3.2, it is once again noted that comets of the Jupiter family (such as 73P; Fig. 2b) and other comets (such as C/1999 S4; Fig. 5b) behave similarly.

4. NATURE OF EPISODIC AGING OF COMETS

The three modes of episodic aging (Secs. 3.1–3.3) have one major common point, as they all involve objects of similar masses, up to $10^{14}$ g, which is a substantial fraction of the mass of a smaller periodic comet, say, 1–2 km in diameter. The modes described in Secs. 3.1 and 3.3 also involve objects of the same inferred shape — highly flattened or pancake-like. In an effort to determine their nature on the comet’s surface, prior to their release, it is desirable to search for a hypothesis that would simultaneously be consistent with all three modes of episodic aging. The mass, and possibly the shape, should be the physical properties to guide one in this search.

4.1 Surface Morphology of Cometary Nuclei from Close-up Imaging Onboard Dedicated Missions

Referring to the Deep Impact images of the nucleus of comet 9P/Tempel, Thomas et al. emphasized that a major aspect of the comet’s geology is strong evidence of layering. While the authors stated that there was suggestive evidence for layering on comets 19P/Borrelly and 81P/Wild, visited, respectively, by the Deep Space 1 and Stardust spacecraft, only in the images of 9P/Tempel was layering widespread and obvious. Thomas et al. identified at least three layers 50 to 200 meters thick and several thinner layers, 1 to 20 meters thick. The layers are readily seen in an image taken by the Deep Impact spacecraft and reproduced in Fig. 6.

Belton et al. proposed a hypothesis that this layering is ubiquitous on nuclei of Jupiter family comets and is an essential property of their internal structure. The nuclei of these comets would then be primordial remnants of the early agglomeration phase of evolution and their interiors would be little modified since the formation time. Belton et al.’s “layered pile” model consists of an ill-defined core, which represents the original aggregate that initiated the growth of the nucleus, and an overlaid pile of randomly stacked layers that date back to the first million years of the accumulation period. The boundaries of the layers are marked by scarps that show the superposition of one layer upon another. While Belton et al. admit that the layers have a wide range of base-area dimensions and thickness, their typical layer covers a base area of $\sim 5$ km² and is $\sim 50$ m thick. With their assumed bulk density of 0.4 g/cm³, the typical layer’s mass is $\sim 10^{14}$ g, on the same order of magnitude as first-generation fragments of split comets (Sec. 3.1). The typical layer’s maximum-to-minimum dimension ratio of $\sim 50$ is reasonably consistent with pancake-like shape (Sec. 3.1).

Could thick layers be identical with major fragments (and, similarly, thin layers with minor fragments) of split comets? Could the layers be lifted off from the nucleus and jetted into the atmosphere? Would they survive the transportation intact or at least not too severely damaged? And what mechanism would be necessary to achieve all this? And since it is known that Jupiter family comets and comets of other categories (such as Halley-like, Oort-cloud, etc.) show no fundamental differences in their behavior after splitting, how important is in this context the fact that the “layered pile” model applies only to Jupiter family comets? Could other kinds of layers be jettisoned from comets of other categories?
4.2 Possible Fragmentation Mechanism

While further research should test the veracity of the layered pile model, the absence of talus deposits along scarps of layers suggests that debris is being swept away by sublimation, perhaps from “below.” This evidence, as well as the difficulty to identify specific surface landforms as vents, is favorable to the possible existence of isolated reservoirs of subsurface ice (primarily water ice) that separate the layers from one another and help holding them together. If layers should be primordial, so should interlayer sheets of ice, in which case water ice is in amorphous phase.

Only after a comet was perturbed into a heliocentric orbit smaller than Jupiter’s, was a topmost layer penetrated and an underlaid ice reservoir reached by the front of a thermal wave, spreading from the surface into the interior and capable to warm up a small fraction of the sheet’s volume to 140–160 K. At that time, an exothermic reaction was set off by a transition of water ice from amorphous phase to cubic phase and began to release energy at a rate of $1 \times 10^6$ erg/g, rapidly heating up the rest of the icy sheet to a temperature exceeding 200 K. The crystallized ice sublimated rapidly in the cavity and the pressure increased by orders of magnitude, reaching several pascals. The molecular drag on the base of the overlaid layer exceeded the comet’s gravity and began to lift off the layer. The separated pancake-shaped fragment started its journey through the comet’s atmosphere, also assisted by rotation, especially if the spin rate was high.

In a different context, this mechanism was more than 30 years ago proposed as an energy source for comet outbursts. The amount of information on amorphous water ice has recently grown exponentially, and it is impossible to present even an abridged overview. From the astrophysical standpoint, laboratory work have provided especially helpful insights into the complexities of the water-ice transitions.

At sublimation pressures of several pascals, the molecular drag is clearly lower than the mechanical strength derived for cometary nuclei, so one would expect this jettisoned first-generation fragment to survive the liftoff, with only a fraction of its mass along the edges disintegrating into dust particles, pebbles, etc. This debris will appear first as an outburst and may later contribute, together with the fragment’s own activity, to its coma and tail. This scenario fit the case of a split comet with one companion. If the layer breaks into several major pieces upon the liftoff, one has the case of a multiply split comet, with the same separation time for all companions.Obviously, there could be other variations of, and complications to, this scenario, such as a layer lifting off in parts, several fragments at different separation times, early signs of cascading fragmentation of the first-generation fragment, etc.
The mechanical strength is likely to vary from comet to comet as well as from one part of the nuclear surface to another. Thus, on rare occasions, the local terrain may be cemented so poorly that the layer fails to survive the liftoff. Instead, the mass immediately collapses and begins to crumble precipitously into a wide cone of mostly microscopic (micron- and submicron-sized) particles that are rapidly accelerated by the expanding gas flow to subkilometer-per-second velocities, soon to make up a steadily growing, sharply-bounded halo. This is the case of a Holmes-like exploding comet (Sec. 3.2), which thus becomes yet another variation of the same fragmentation process.

If the comet is reduced to a few or just a single layer or if there is no core, such as apparently may be the case with C/1999 S4, the liftoff phase is largely irrelevant. It is only a matter of time before the process of cascading fragmentation sets in, resulting in the comet’s cataclysmic disintegration (Sec. 3.3). Thus, the proposed conceptual mechanism explains the whole range of events of fragmentation-based episodic aging of comets. But because both Jupiter family comets and comets of other origin are considered, there is only a loose relationship between this set of scenarios and the original layered pile hypothesis.

4.3 End States

In or out of the context of the layered pile model, the end state of an aged comet depends on the original body situated in its center, regardless of the pile of layers. If the core is structurally strong enough to survive the process of removal of all individual layers, one by one, the comet is eventually reduced to this core. Depending on its physical properties, it may still stay a little active for a while or become an inert asteroidal object right away. If there is no core — the case that apparently includes cataclysmically fragmenting comets — there is left nothing but scattered dust and the comet vanishes.

5. CONCLUSIONS

Comets age by shedding mass and their nuclei grow smaller and less massive with time. Normal activity of comets, consisting of sublimation of water ice and more volatile substances together with expulsion of refractory material, is episodic because it is a sum of contributions from individual discrete emission sources on or below the surface of the nucleus, which are subjected to the processes of activation, deactivation, dormancy, reactivation, and extinction due to their ice reservoir’s ultimate depletion. Dynamical techniques have proven most useful to our efforts of understanding the influence of these processes on the ever changing activity of comets, especially thanks to effects of discontinuous orbital anomalies that are experienced by the motions of many periodic comets. Besides the intrinsic properties of each emission source, cometary activity is subjected to diurnal and seasonal variations, which depend on the rotation and precession (if any) of the nucleus.

Less frequent, but with greater indelible imprints, are events of episodic aging that involve fragmentation of the nucleus (Sec. 3.1), Holmes-like explosions (Sec. 3.2), and, especially, cataclysmic disintegration (Sec. 3.3). These events represent major steps in the aging process, but their temporal rate is not well known. The proposed hypothesis is based on a layered pile model originally advanced for Jupiter family comets, but expanded here to include all comets. The idea is that as the layers peel off, the nucleus diminishes in size, a sign of comet aging. It is postulated that the fragmentation process is set off by an exothermic reaction caused by a transition of water ice (warmed gradually up to 140–160 K by a thermal wave penetrating from the surface into the interior of the nucleus) from amorphous phase to cubic phase in a reservoir spread below a layer of up to $10^{14}$ g in mass. The jetisoned layer becomes a free-flying pancake that either survives as a companion nucleus to the principal fragment, disintegrating only gradually in space, or it collapses upon the liftoff and is, in a wide cone, scattered as a cloud of fine dust into the atmosphere to make up a rapidly-expanding, sharply-bounded halo, typical for Holmes-like exploding comets. These explosions, events of nearly global proportions on the scale of a comet’s nucleus, should not be confused with ordinary outbursts that originate from small isolated emission sources and carry a mass of dust that is many orders of magnitude lower.

The two categories of episodic aging (sublimation and dust emission versus nuclear fragmentation) both contribute to the rate at which comets grow older, but fragmentation can dramatically accelerate the pace. The end state of an aged comet is either a small inert asteroid-like object or nothing more than a rapidly vanishing cloud of scattered dust.

ACKNOWLEDGMENTS

I thank C. W. Hergenrother and P. Vasey for permission to reproduce their images of comets. This work was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.
REFERENCES