Origin of Comets

The Life, Times and Persistent Puzzles of Comets

Mark E. Bailey
Armagh Observatory

http://star.arm.ac.uk/ meb@arm.ac.uk
http://climate.arm.ac.uk

Some Famous Comets — I

Comet C/1858 L1 (Donati)

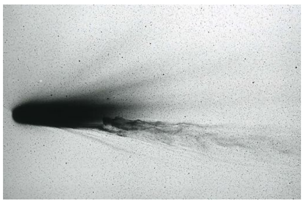
- Discovered 2nd June 1858.
- Brightened through July and August.
- Easy naked-eye object during September and October that year.

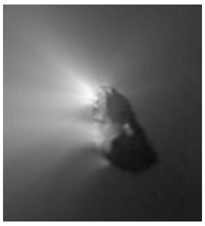
Described by many as "The most beautiful comet of all time!"

Some Famous Comets — II

Comet 1P/Halley)

- Perhaps the most famous periodic comet.
- Returns every 75–76 years.
- ROE/UKST image (top) shows great tail 'disconnection' event of 1986 March 9.
- Nucleus imaged by ESA Giotto spacecraft 1986 March 14 (H.U. Keller)
  - size ∼15.3 ×7.2 × 7.2 km
  - average albedo ∼0.04
  - only 10–20% of surface 'active'





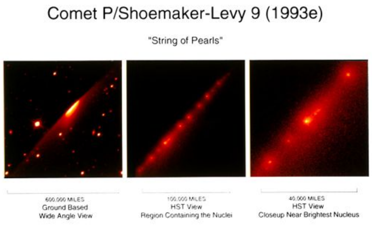
Some Famous Comets — III

- Discovered 1995 July 23 by Alan Hale and Thomas Bopp.
- A 'great comet', the best many of us will remember.
- Visible for several months during Spring 1997.
- Image signed by Thomas Bopp 1997 June 20, taken on 1997 March 28.

Comet C/1995 O1 (Hale-Bopp)

Some Famous Comets — IV Shoemaker-Levy 9

- Comet D/1993 F2 Discovered 1993 March 25.
- Previously passed within Roche limit of Jupiter on 1992 July 8; broke into fragments.
- These fragments (the observed SL 9 comet) impacted on Jupiter from 1994 July 16–22.
- Impacts and impact scars visible from Earth.



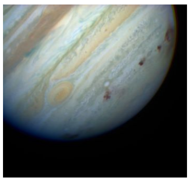
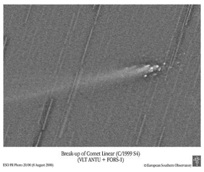


Image credits: H. Weaver & T. Smith; NASA/ESA

Various Cometary End-States

Disintegration



Break-up of Comet C/1999 S4 (LINEAR), 2000 August. Image credit ESO.

Outgassing

Comet C/1996 B2 (Hyakutake), 1996 March. Image credit D. Diereck.

Sun-Grazer

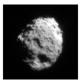
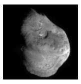
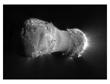
58. Comet Ikeya-Seki, 1965. Perihelion passage October, 21·1831 U.T. Photograph by Alan McClure, Mount Pinos, California, taken with a 10 in. f/3·8 lens, panchromatic plate, 1965, November 1, 12 h. 43 m.–12 h. 54 m.

Comet C/1965 S1 (Ikeya-Seki), 1965 October. Image credit A. McClure.

Unresolved Questions

What *are* comets, and why so diverse?
How are they formed, and where?
Where do they primarily come from now?
What effects do they have on Earth (and Sun)?
How do they die and where do they go?

- dynamical ejection from solar system;
- collision with planets, or with Sun;
- evolution to inert end-state: e.g. by outgassing or formation of inert crust;
- physical decay and disintegration: e.g. loss of volatiles and dust, splitting, breakup etc.

Comet:	81P/Wild 2	19P/Borrelly	9P/Tempel 1	103P/Hartley 2
Size:	∼4.2 km	∼5.0 km	∼6.1 km	∼1.4 km
				

Credits: NASA/JPL/Stardust/Deep Space 1/Deep Impact

Foundations of Astronomy

Two main themes:

1. Advances in general science and understanding that began in the Renaissance, i.e. the few centuries up to the Industrial Revolution;
   - Provide the backdrop against with to 'read' the historical literature on comets in the 17th–19th century and earlier;
   - End of the 18th century becomes a kind of 'watershed' between an older pre-scientific view of the natural world, and more modern 'scientific' views.
2. A more or less continuous strand of interest in comets and cometary debris, from earliest times right up to the present, viz:
   - The physical and societal impact of comets;
   - Comets as agents of destruction (catastrophism) versus celestial bodies that convey(ed) ingredients necessary to sustain 'Life' on Earth;
   - The rejection and modern rediscovery of cometary catastrophism (aka impact hazard).

New paradigm: Earth in touch with its near-space environment.

Summary Up to 18th-Century Watershed

1. Comets sometimes the most prominent objects in sky: as bright as the brightest stars. Appear unpredictably but move like planets . . .
2. Comets and associated meteoric phenomena scrutinized as 'omens' to predict events on Earth; observations go back millennia.
3. Mankind's puzzling 'fear' of comets still not adequately explained. Historical evidence suggests 'the sky' may have been significantly different in proto-historic times.
    - Suggests actual experience of Earth as an 'open' system: in touch with its near-space celestial environment.
    - But is astronomical change on such historical timescales possible?
4. Two thousand years of 'atmospheric' ideas finally superseded by a correct 'celestial' picture for comets.
5. Later — and modern — debates focus on 'interstellar' versus 'solar system' ideas, and whether comets are distinct objects in their own right, or with an origin more akin to planets.

Early 'Planetary' vs 'Interstellar' Debates

Increasing database of comets — new observations plus translations of historic (mostly Chinese) records — leads to comets established as definitely celestial. But are they solar system or interstellar in origin?

1. Most comets have parabolic orbits — quite different in shape and inclination from the planets
2. The few 'periodic' orbits (e.g. 1P/Halley) could be — and were — open to question, e.g.
    - how far did Newton's new law of gravity ('action at a distance') extend
    - how sure can we be that two comets with similar orbital elements are the same comet, or — coincidentally — two different comets?
3. Return of Halley's comet in 1759, as predicted, a significant advance
4. But Lexell's comet (1770) proved the matter was not clear-cut
    - The observed orbit was definitely elliptical, but the comet had previously been 'captured' and was subsequently 'ejected' by Jupiter!

Lexell's Comet: Complex Dynamical Evolution

1. Discovered by Messier on 14/15 June 1770; soon recognized as having an unusually large daily motion;
2. Exceptionally close approach to Earth ($\Delta_{\mathrm{E}} = 00146\,\mathrm{AU}$) on 1st July 1770, the closest approach of any comet to Earth in recent times;
3. First comet to have mass estimated ($<1/5000\,M_{\oplus}$);
4. First comet to have elliptical orbit determined from observations of just one revolution, i.e. $a \simeq 3.15\,\mathrm{AU}$, $q \simeq 0.67\,\mathrm{AU}$, $P = 5.6\,\mathrm{yr} \simeq 0.5 P_{\mathrm{J}}$ — but comet never seen again!
5. Later calculations show comet had an original orbit with $q \simeq 2.9\,\mathrm{AU}$ and $P \simeq 9.2\,\mathrm{yr}$. Exceptional close approach to Jupiter ($\Delta_{\mathrm{J}} = 0.02\,AU$, on 27 March 1767) led to observed Earth-crossing orbit destined to pass just 2 Jovian radii above Jupiter ($\Delta_{\mathrm{J}} = 0.002\,AU$, on 2 July 1779), which drastically changed orbit again;
6. New orbit has perihelion near Jupiter ($q \simeq 5.17\,\mathrm{AU}$) and semi-major axis $a \simeq 43.0\,\mathrm{AU}$, i.e. $P \simeq 280\,\mathrm{yr}$. Illustrates complexity of cometary dynamical evolution.

Laplace's 'Nebular Hypothesis': Dominating 19th Century

Laplace, the French 'Newton', an intellectual giant. Comets are gravitating, celestial bodies; but are they 'solar system' or 'interstellar' in origin?

1. Laplace's Nebular Hypothesis (1805). Independent of Kant's (1755) nebular theory; emphasizes the regularities of the planetary system, e.g. the central dominance of the Sun, the planetary orbits being nearly circular, coplanar etc. against the irregularities of the comets.
2. Provided a physical model based on Newtonian principles and the collapse and contraction of an initially slowly rotating promordial gas cloud.
    - Contraction leads to 'spin-up' and shedding of rings of gas from the central body's equator $\Longrightarrow$ formation of the planets;
    - Most of the mass remains at the centre; $\Longrightarrow$ formation of Sun
3. But comets remain anomalous; therefore must have a separate, i.e. 'interstellar' origin.

NB: Kant (1755) had developed a different nebular theory and believed that comets *could* be incorporated into such a model, but Laplace was in the ascendant.

Formation of Solar System Following Laplace

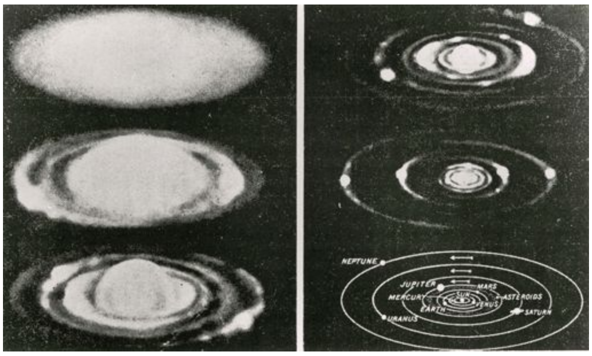


Illustration of origin of solar system according to Laplace's Nebular Hypothesis, after Whipple (1964)

Lagrange's 'Planetary' Theory for Comets

1. Emphasized discovery of first asteroids: Ceres and Pallas, by Piazzi and Olbers in 1801 and 1802 respectively
    - Ceres originally described by Piazzi as a 'star-like' comet;
    - Early suggestions of transient nebulosity around both objects (Herschel, Schröter); confirmation elusive;
2. The new objects ('asteroids' or minor planets) upset underlying simplicity of Laplace's nebular hypothesis, proving there exist planetary objects in relatively high-inclination, non-circular orbits.
    - Olbers speculates they might be fragments of a former much larger planet between Mars and Jupiter — blown to pieces by internal forces, or the impact of a comet . . .
3. Lagrange (1814) posthumously publishes his proposal that at least some comets might have a similar origin.
4. Thus a new 'solar system' theory for comets is born, harking back to earlier 'planetary' and catastrophic notions about comets.

Laplace (1805) versus Lagrange (1814)

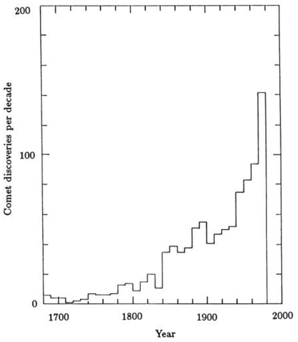
1. Laplace argues that his Nebular Hypothesis explains the main regularities of solar system, but not the comets; and that:
    - if Sun at rest with respect to interstellar population of comets, then we would expect to see an overwhelming excess of near-parabolic orbits — as observed;
    - anomalies, such as Halley's comet, can be explained by appeal to *ad hoc* planetary 'capture' mechanism, as per Comet Lexell.
2. Lagrange and others instead focus on growing number of known 'short-period' comets:
    - Too many to be explained by *"ad hoc"* chance capture, but formation of comets remains unexplained on the planetary collision/explosion theory, therefore equally *ad hoc*.

Neither theory perfect: Laplace's founded on an assumption about solar motion (later proved wrong); Lagrange's on an assumption about the (catastrophic) origin of comets.

Either way, the near-parabolic excess remains problematic (perhaps there are two types of comet, e.g. long-period vs. short-period?).

New Discoveries c.1800–1900

1. Rapid increase in # of known comets; mostly parabolic.
2. Whereas in 1790 there were 78 parabolic comets (including 7 sightings of Halley's comet); by 1890 there were c.270, of which 20 were definitely periodic and 43 definitely elliptical.
3. $\Longrightarrow$ growing focus on the 'periodic' subsample: a group with mostly low-inclination, 'direct' orbits and greater orbital similarity to asteroids than to other comets.
4. Suggests a 'planetary' solar-system source for these 'short-period' comets.



Increasing rate of comet discoveries with 'known' orbits, 1680–1980

A Key Observation: The Solar Motion

1. Interstellar theory depended on Sun essentially at rest with respect to assumed interstellar population;
2. Herschel (1783) previously concluded Sun was moving, but results not statistically significant (e.g. Biot, Bessel); Gauss (1815) the first to highlight its potential importance for comets;
3. Argelander (1837) demonstrates reality of solar motion; but by then its implications for comets 'forgotten'; Laplace's 'interstellar' theory survives.
4. Carrington (1860, 1863) emphasizes that solar motion would produce more comets coming from one direction than another (cf. abberation). But cannot see any effect — and concludes solar motion is wrong!
5. Peirce (1849) and Schiaparelli (1860) put argument right way around, but their work not accepted: e.g. Newton (1878): *"Professor Schiaparelli by introducing (improperly, as I am sure he will concede) the motion of the Sun in space was led to decide against a foreign origin for comets."*
6. Fabry (1893) finally overthrows Laplacian paradigm: concludes Sun surrounded by comoving comet cloud; hence comets are 'Solar System' objects

---

A 'New' Solar System Paradigm c.1900: Laplace Rejected

1. Number of short-period comets a growing problem for 'capture' hypothesis.
2. If short-period comets are 'solar system' objects, perhaps they formed — like asteroids — in a 'planetary disruption' event (cf. Lagrange 1814, Proctor 1870); or by break-up of a single, large comet (Alexander 1850, Bredichin 1889) ...
    - In any case, short-period comets seem likely to be 'solar system' in origin; and so too — somehow — must those with long-period 'parabolic' orbits.
3. Laplace's 'interstellar' theory therefore rejected; leads to a return to Kant's 'solar nebula' picture.
4. Two new themes at this time: (1) if comets are 'solar system' in origin, then what do comets tell us about origin of solar system?; and (2) observed orbits are probably not those that originally prevailed.

Theories now suggest observed orbits are example of Darwinian evolution. Parabolic excess explained by Darwinian 'survival of the fittest' — from an initial primordial population containing all possible orbits.

Pathways to the Oort Cloud c.1920–1950

1. New age for Earth around 1920 (billions of years rather than tens or hundreds of millions) exacerbates survival problem for parabolic comets.
    - The 'Darwinian' argument implies we should see no comets at all!
    - And the 'capture' theory still fails to explain number of short-period comets.
2. Most astronomers remain irrationally attached to 'solar system' models.
3. Nevertheless, some seek to develop alternative 'interstellar' theories, e.g. (1) recent 'capture' of comets by passage of Sun through a dense 'interstellar cloud' of comets (Nölke 1909, Bobrovnikoff 1929); and (2) comets formed by passage of Sun through a dense interstellar dust cloud (Lyttleton 1948).
4. Öpik (1932) notes Darwinian argument depends on comets always being subject to decay, and develops first quantitative theory of stellar perturbations on long-period orbits.
    - He shows cometary perihelia evolve outwards (and inwards) under stellar action; the key ingredient for the 'Oort cloud' picture.

Summary of 19th and Early 20th-Century Theories

Originator and Date	Theory and Broad Type
Laplace (c.1805)	Comets are interstellar objects
Lagrange/Brewster (c.1810)	Comets are result of planetary explosion
Peirce (1849)	Comets are solar system objects, originating at great heliocentric distances
Schiaparelli (c.1860)	Comets are solar system bodies, occupying a huge co-moving swarm about the Sun
Proctor (1884)	Comets are formed in vast planetary explosions
Newcomb (c.1910)	Comets originate in the collapsing protosolar nebula
Chamberlin/Crommelin (c.1920)	Short-period comets originate in planetary explosions or ejection of material from Sun
Bobrovnikoff (1929)	Comets are interstellar; occur in dense interstellar swarms and have been recently captured
Vsekhsvyatski (c.1930–1980)	Comets originate in explosions on the outer planets or their satellites
Öpik (1932)	Comets are primordial solar system bodies; long-period orbits affected by stellar perturbations
Nölke (1936)	Comets are interstellar; recently captured via interaction with dense resisting medium

---

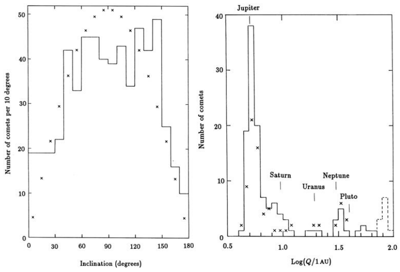
Key Arguments 1920–1950: (1) Orbit Statistics

LPcomets ($P > 200$ yr): (i) nearly isotropic; (ii) strong 'parabolic excess'.

Range $1/a$	N
0.000–0.002	177.0
0.002–0.004	10.0
0.004–0.006	8.0
0.006–0.008	7.0
0.008–0.010	2.5
0.010–0.012	6.5
0.012–0.014	1.0
> 0.014	9.0

LP comets (1850–1936) $1/a$ distribution (Van Woerkom 1948).

SP comets ($P < 200$ yr): (i) most low-inclination; (ii) aphelion distances correlate with planets



Left: Near-isotropic distribution of LP cometary inclinations (589 LP comets up to 1982).

Right: Distribution of SP comet aphelia Q correlating with planetary semi-major axes (121 periodic comets up to 1982).

Key Arguments 1920–1950: (2) Planetary Perturbations

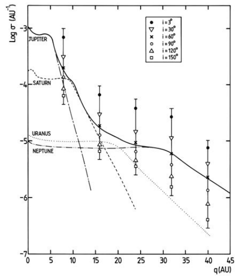
1. Observed orbits cannot be the 'original' orbits.
2. Planetary perturbations mostly change the orbital energy E per unit mass, i.e. lead to $\Delta E = \Delta(-GM_{\odot}/2a) \propto \Delta(1/a)$.
3. For randomly distributed close encounters (i.e. strong perturbations), the chance of an energy change E is $\propto |E|^{-3}$, i.e. (for encounters with Jupiter):
$$\phi(E)\,dE \simeq \frac{8}{3}\left(\frac{M_J}{M_{\odot}}\right)^2 \frac{1}{a_J^2}|E|^{-3}\,dE$$
4. Leads to concept of 'capture probability', p_c (cf. Lexell), i.e. probability to evolve in one revolution to semi-major axis less than a:
$$p_c(<a) \simeq \frac{4}{3}\left(\frac{M_J}{M_{\odot}}\right)^2 \left(\frac{a}{a_J}\right)^2$$

$\Longrightarrow$ Capture probability to period $P < 200$ yr is $\lesssim 5 \times 10^{-5}$, i.e. small; emphasizes problem of number of observed short-period orbits.

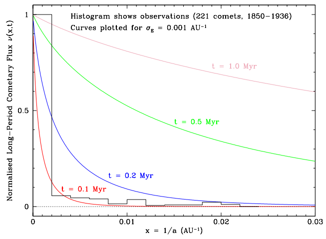
Key Arguments 1920–1950: (3) $1/a$-distribution and Diffusion Theory

1. Changes in $x = 1/a$ akin to random walk in E, i.e. energies diffuse.
2. Define σ_E = r.m.s. $\Delta(1/a)$ per revolution; $\sigma_E \lesssim 10^{-3}\,\mathrm{AU}^{-1}$.
$\Longrightarrow$ after N revolutions, expect $x \approx \sqrt{N}\sigma_E$, i.e. $\approx 10^3$ revolutions to reach $P < 200$ yr, i.e. evolution by many small perturbations.
3. Solve diffusion equation (ignoring decay):
$$\frac{\partial \nu}{\partial t} = \frac{\sigma_E^2}{2P(x)}\frac{\partial^2 \nu}{\partial x^2} \qquad \text{where} \quad P(a) = 2\pi (GM_{\odot})^{-1/2} a^{3/2}$$
4. Leads to:
$$\nu(x,t) = \left(1 + \frac{t_D}{t}\right)\exp(-t_D/t) \qquad \text{where} \quad t_D = 8\sqrt{x}P(x)/\sigma_E^2;$$
i.e. diffusion time-scale $\lesssim 1$ Myr.
5. $\Longrightarrow$ rapid evolution of $1/a$-distribution to a quasi-steady state ...

Solution of Diffusion Equation



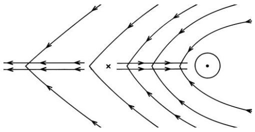
σ_E distribution versus q (AU^{-1}) for randomly distributed initial parabolic orbits. Very roughly, $\sigma_E(q) \approx 10^{-3}\exp(-q/5.2\,\mathrm{AU})\,\mathrm{AU}^{-1}$.



Solution of diffusion equation with no decay (after Van Woerkom 1948); note rapid evolution to 'flat' distribution, completely contrary to observed $1/a$-distribution (histogram).

Lyttleton's (1948) Accretion Theory

1. A novel variant of the interstellar hypothesis; the first to address both where comets come from and how they are formed.
2. Consider motion of Sun through a dense dust cloud of density ρ_{dust}. Collisions of dust grains on axis of symmetry dissipate energy and cause some grains to be captured — these coalesce to become proto-comets.



3. Get inflow within a stagnation radius r_0, approximately the accretion radius $R_A = GM_\odot/V^2$. For $V = 5\,\text{km s}^{-1}$, $R_A \simeq 35\,\text{AU}$.
4. In a steady-state, the stream mass per unit length is $\mu \simeq 2\pi\rho_{\text{dust}}R_A^2$ and the stream velocity V_s is roughly the free-fall speed from R_A. Thus, any new comets have initial semi-major axes $a \lesssim R_A/2$.

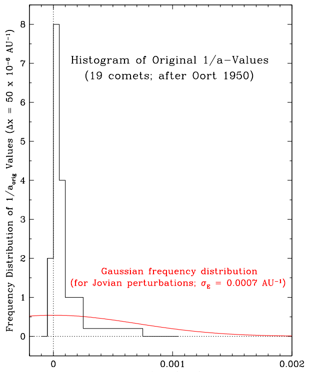
Problems with Lyttleton's Theory

1. Dust clouds do not exist on their own. The interstellar dust is dominated (by a mass fraction of at least a factor of 50) by hydrogen gas. Effects of gas must be included; this was never done.
2. The supposed proto-comets are far too small. Even if an accretion stream could be set up, only very short segments of length $d(r) \lesssim 2\sqrt{\mu r^3/M_\odot}$ at heliocentric distance r could successfully contract against the tidal field of the Sun. This leads to $m_c \lesssim 10^8(10\,\text{km s}^{-1}/V)^9(\rho_{\text{dust}}/10^{-22}\,\text{kg m}^{-3})^{3/2}\,\text{kg}$.
3. Inital orbits too short period and too anisotropic. Lyttleton argues for a long period of randomisation of orbits following the last accretion episode, but then the predicted $1/a$-distribution quite wrong (diffusion theory).
4. The supposed proto-comets are on initial orbits directed towards Sun (or solar-system barycentre). All the initial comets will fall onto the Sun, unless inhomogeneities or planetary perturbations are invoked to deflect the stream.

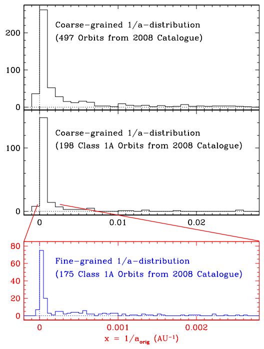
In summary, despite strong advocacy of theory by Lyttleton for next 30 years: "The theory is disproved: an honourable fate for a good theory"!

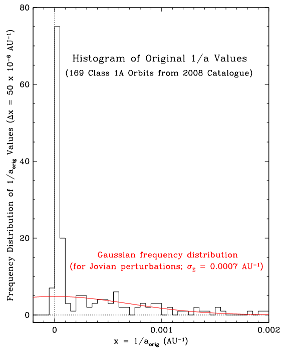
Birth of a Theory: The 1950 Oort Cloud

1. Oort considers the original $1/a$-values of the 19 most accurate orbits; i.e. those with mean errors $< 10^{-4}\,\text{AU}^{-1}$.
2. Enables fine-grained binning of $1/a$-distribution for first time.
3. More than half had 'original' $1/a$-values $< 50 \times 10^{-6}\,\text{AU}^{-1}$; and none had $1/a > 750 \times 10^{-6}\,\text{AU}^{-1}$.
4. Note the extreme narrowness of the sharp peak in the distribution of 'observed' original $1/a$-values.



Comparison with Modern Data





Argument for Oort Cloud — c.1950

1. Sharp spike in observed $1/a$-distribution rules out interstellar capture (Van Woerkom 1948); and Lyttleton's 'capture theory' (1948) seriously deficient …suggests comets have primordial solar system origin, and the observed comets are coming into inner planetary region for the first time.
2. If comets are primordial, there must be a 'comet store' — the 'home' of the comet — somewhere beyond the zone of visibility, where comets can survive. Logically, this must contain comets in orbits of large perihelion distance.
3. Oort then addresses how to get comets from safe storage into inner solar system:
    - Planetary perturbations? — NO: they broaden the $1/a$-distribution too much (van Woerkom), contradicting observations.
    - Resistance of dense interstellar medium? — NO: it is implausible, and such a medium would primarily affect the comets' aphelion distances, again contradicting observed $1/a$-distribution.
    - Stellar perturbations? — YES: cometary orbits extend up to halfway to the nearest star; they must be affected by passing stars (cf. Öpik 1932).

Argument for Oort Cloud — in Modern Terms

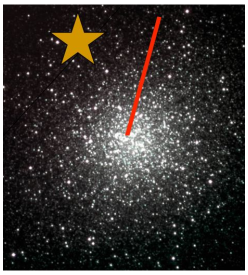
1. Observations $\Longrightarrow$ We see ~ 1 'new' comet ($q < 5\,\text{AU}$, $H_{10} < 7$) discovered per year.
    - Semi-major axes $a > 2 \times 10^4\,\text{AU}$: i.e. near parabolic limit; orbital periods $P \simeq$ 3–30 Myr — short compared to age of solar system.
    - These so-called 'new' comets strongly perturbed by Jupiter, so that roughly half ejected, the remainder 'captured'.
    - 'Captured' comets return, to be ejected or lost to short-period orbits and eventual decay.
2. Conclude: All observed comets are ultimately lost; and the 'loss cone' affects all orbits with $q \lesssim 15\,\text{AU}$. The loss timescale $\lll$ age of solar system.
3. $\Longrightarrow$ comets are either a transient phenomenon, or there is a long-lived reservoir to replenish those that are lost.
4. Oort adopts primordial 'steady-state' hypothesis.

Further Details — in Modern Terms

1. 'New' comets are only lost if q lies within loss cone, i.e. $q \lesssim 15\,\text{AU}$; $\Longrightarrow$ Oort's reservoir must contain long-period comets of large q.
2. For long-period orbits, planets change the orbital energy, i.e. change $1/a$, keeping q nearly constant; stars change the angular momentum, i.e. change q, keeping $1/a$ constant.
3. The change in q is about the size of the loss cone, provided the orbit is large enough.
    - Δq per revolution $\propto a^{7/2}$, i.e. depends sensitively on a.
    - $\Longrightarrow$ the reservoir must contain orbits of very long period ($a > 2 \times 10^4\,\text{AU}$, $P > 5\,\text{Myr}$) — just like the observations.
    - Leads to Oort's idea of a nearly spherical cloud of comets with orbits extending up to halfway to nearest star.
4. The cloud is 'gardened' by various external perturbations.
    - including stellar, molecular cloud and large-scale systematic effects of Galactic tide.

View of Oort Cloud

1. Like a globular star cluster, such as M13…
   Imagine Sun at centre
    - the stars become 'comets'
    - the shape (like a flattened rugby ball) is about right
    - the strong concentration of comets towards the centre is about right
    - the overall dynamics is similar
2. Can calculate 'families' of Oort cloud models, in the same way as for star clusters and galaxies
3. External perturbations (e.g. stars) change cometary orbits



The 'loss cone' behaves just like the loss cone around a black hole in a galactic nucleus

Standard (1950) Model

1. Assume: (1) spherical symmetry; outer radius $R_0 \simeq 200,000\,\mathrm{AU}$; (2) random velocities, 'gardened' by stellar perturbations; (3) hydrostatic equilibrium (cloud neither expanding nor contracting); and (4) a simple energy distribution, e.g. a power-law distribution of orbital energies per unit mass $E = -GM_\odot/2a$.
2. If $f(E)\,dE \propto |E|^{-\gamma}\,dE$, then the number density $n(r)$ is roughly proportional to $r^{\gamma-4}$.
3. Oort's (1950) model has $\gamma = 5/2$, corresponding to velocity space being uniformly filled at r up to a value $V_{\max}$ equal to the free-fall speed from R_0 to r. This implies $n(r) \underset{\sim}{\propto} r^{-3/2}$, i.e. most of the mass near the outer edge.
4. Other models have smaller γ (e.g. $\gamma \sim 0$), and much sharper inward density increases. The structure is much more like a dense star cluster, with a strong concentration of mass towards the centre, not a shell.
5. Leads to the concept of an inner Oort cloud, i.e. a Dense Inner Core: a region inaccessible to observation but possibly containing most of the cometary mass, and relatively safe from external perturbations.

Structure of Oort Cloud

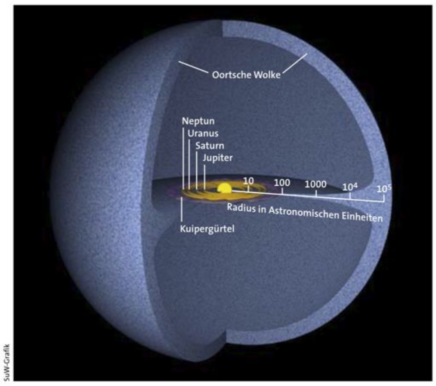


Image © Sterne und Weltraum (2011 Feb., p.20).

Summary of 20th-Century Theories: c.1940–1970

Originator and Date	Theory and Broad Type
Edgeworth (1938–1961)	Comets are solar system objects formed by accretion in outer protoplanetary disc beyond Neptune
Lyttleton (1948)	Comets are swarms of interstellar dust grains, concentrated and captured into solar system by the accretion mechanism
Oort (1950)	Comets originate in asteroid belt or in disruption of a former asteroidal planet; now in a vast swarm about the Sun
Kuiper (1951)	Comets accumulated from icy condensates in or beyond Neptune-Pluto zone of low-mass protoplanetary disc
Krat (1952)	Comets formed by collisions of bodies beyond Neptune
Katasseff (1955)	Comets formed by disruption of a former asteroidal planet
Cameron (1962)	Comets formed by accretion beyond Pluto in massive, extended protoplanetary disc
Donn (1963)	Comets are fractal aggregates of interstellar grains
Whipple (1964)	Comets formed in outer region of protoplanetary disc, either in Uranus-Neptune-Pluto zone or beyond Pluto
Öpik (1966–1978)	Comets accumulated in Jupiter-Saturn region of a low-mass protoplanetary disc
Safronov (1967–1978)	Comets formed in Uranus-Neptune zone of low-mass protoplanetary disc

Summary of 20th-Century Theories: c.1970–1978

Originator and Date	Theory and Broad Type
Alfvén & Arrhenius (1970)	Comets formed by accretion of dust in 'jet streams' in early solar system; on-going formation of short-period comets
Hills (1973)	Comets accumulated from ice-covered dust grains in Uranus-Neptune region of a low-mass protoplanetary disc
Donn (1973, 1976)	Accumulation of comets in separate fragments of Sun's collapsing parent interstellar cloud
Reeves (1974)	Comets form continuously near the heliopause
Mendis & Alfvén (1973–1976)	Comets formed or continue to form in jet streams
McCrea (1975)	Comets formed by gravitational instability involving ice-covered dust grains within a quiescent interstellar cloud
Van Flandern (1975–1987)	Comets formed in *recent* disruption of a former asteroidal planet, c.5 Myr ago
Whipple & Lecar (1976)	Comets formed in interaction of solar wind with Sun's surrounding circumstellar nebula
Woolfson (1978)	Comets formed by collisions between planets or protoplanets, in 'tidal' theory for origin of solar system
Drobyshevski (1978)	Comets are ejecta from collisions of moon-sized, ice-covered planetoids beyond Pluto

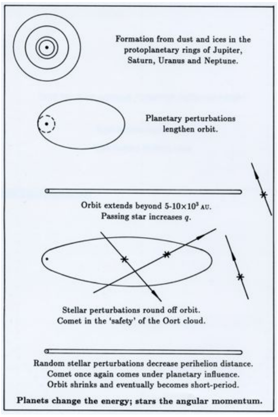
Summary of 20th-Century Theories: c.1978–1990

Originator and Date	Theory and Broad Type
Cameron (1978)	Comets formed in outer region of a massive protoplanetary disc, far beyond Pluto
Biermann & Michel (1978)	Comets formed in outer region of the protoplanetary disc
Biermann (1981)	Comets formed in satellite fragments of protosolar nebula
Hills (1981, 1982)	Comets formed by accelerated collapse of dust clumps in protosolar nebula, driven by differential radiation pressure
Humphries (1982, 1986), Napier & Humphries (1982)	Comets formed by accelerated collapse of dust clumps in molecular clouds, driven by photo-desorptive jet-reaction effect
Yabushita (1983)	Comets formed in interstellar dust 'globules' preceded by dust sedimentation
O'Dell (1986)	Comets continuously formed by accretion of interstellar or solar wind ices onto asteroidal dust ejected into long-period orbits by solar radiation pressure
Bailey (1987)	Comets formed by accretion of interstellar dust grains followed by gravitational instability in wind-driven shells around protostars

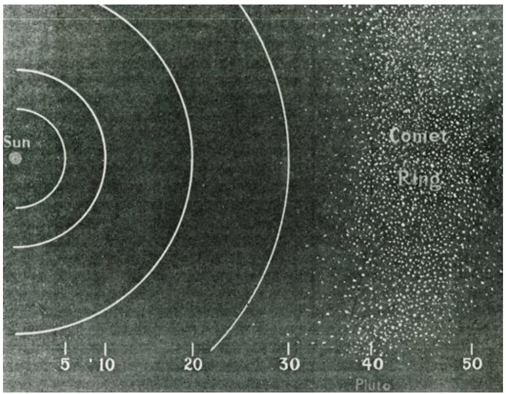
Post-1990 Consensus

1. Comets are primordial solar-system bodies.
2. Formed by coagulation of originally interstellar dust grains during and after formation of protoplanetary disc.
3. Coagulation proceeds rapidly in inner solar system, more slowly farther out (cf. Kant 1755); to produce the protoplanetary building blocks, namely: cometesimals and planetesimals.
4. Late stages of planet formation involve (1) planetary and proto-planet collisions; (2) planetary migration (under mutual gravitational perturbations and evolution of protoplanetary disc); and (3) late-stage bombardment of planetary surfaces by comets and asteroids.
5. Work on origin of comets focuses on (1) dynamical evolution of short-period comets; (2) simulations of origin and evolution of the Oort cloud; (3) origin of Centaurs, in the Jupiter-Neptune region and beyond; and (4) the structure and evolution of the trans-Neptunian region, i.e. the Edgeworth-Kuiper belt and beyond.

Oort Cloud Formation and Evolution Under Planetary Perturbations



Whipple's (1964) Comet Ring: The Prototype EKB



Persistent Puzzles of Comets

1. What are comets . . . Are they all the same . . . Or are some comets (e.g. long-period) different from others (e.g. short-period)?
2. How are comets formed, and where? For example, are comets formed in or beyond the protoplanetary region; or in the protostellar molecular cloud; or beyond, in interstellar space?
3. What is (or are) the proximate source(s) of observed comets?
4. What is the structure and evolution of the 'observed' Oort cloud; how was it formed; and does it contain a massive dense inner core?
5. What is the role of newly discovered outer solar system bodies: Centaurs, Edgeworth-Kuiper belt objects, trans-Neptunian objects etc?
6. What is the cometary mass function and average cometary mass? . . . and how many comets are there, and what is their total mass?
7. Are comets fragile or strong; what are their end-states; and what is the impact of cometary debris on the planets, Earth and Sun?

Steps to Making Comets – I: Pre-Stellar Phase

1. Form dust grains in atmospheres of cool giant stars; eject to interstellar medium (ISM) via stellar winds.
2. Cook in ISM for 10–1,000 Myr: complex cycling of grains through hot diffuse ISM, cool molecular clouds (MCs) and cold MC Cores.
   - In the clouds, grains accrete a frosting of interstellar volatiles; in the hot ISM, ice is sputtered and UV photo-processed; and grains are ground down by collisions and evaporation.
3. Produce interstellar dust with a complex chemistry and broad size distribution; some grains have diameters up to microns or more.
4. Ices on and within the interstellar dust aggregates contain clues to the grains' previous history and to the processes that accompanied their 'final' pre-solar accumulation as part of the Sun's parent molecular cloud.
   - Cometary dust has a rich Cosmic Chemical Memory; cometary dust samples pre-solar history of solar-system material.

Steps to Making Comets – II: Protosolar Disc Phase

1. Form the solar nebula from a rotating protosolar molecular cloud. It cools and collapses to produce a dense gas-and-dust disc.
2. Typical cloud parameters: Temperature $T \approx 10$ K; Radius $R \approx 0.1$ pc; Mass $M \approx 1$–$2\,M_\odot$. Initial disc radius R_d small compared to R, but large compared to current planetary system. For reasonable parameters, $R_d \approx$ few $\times$ 100 AU.
3. Grains grow during nebular collapse and during disc evolution, acquiring a 'frosting' of ices from condensing volatiles in the MC core and protoplanetary disc.
4. In the inner few AU of nebula, dust destroyed by collisions or by heating from the newly formed proto-Sun; dust farther out retains its Cosmic Chemical Memory.

By time Sun forms, expect grains with a complex 'hierarchical' structure, with evidence of both hot (pre-stellar) phases and cold (MC) phases of evolution.

Models of Interstellar Grain Aggregates

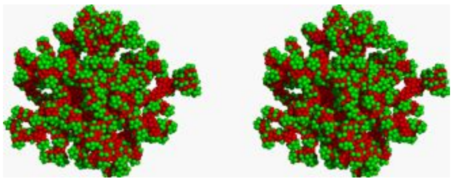


Image credits: Top: © E.L. Wright (UCLA); Bottom J.M. Greenberg (Leiden)

Steps to Making Comets – III: Protoplanetary Disc Phase

1. Condensed ice composition: expect ices such as water, carbon monoxide, carbon dioxide, methanol, hydrogen cyanide, ammonia, methane etc.
2. Disc surface density Σ_d(solids) at 10 AU approximately 10 kg m^{-2}; radial variation roughly a power law, i.e. $\Sigma_d \propto r^{-3/2}$. Gas-to-Dust ratio roughly 50 initially.
   - Surface density corresponds to a traditional 'minimum mass' protoplanetary disc within planetary region; total mass of solids within $\sim$300 AU could range up to several 100 Earth masses.
3. Initial grain growth proceeds rapidly in presence of gas through turbulence-driven coagulation. Large grains initially drift inwards due to gas drag and accrete smaller 'background' grains.
4. Grain radius versus time: $a(t) \approx 0.3\,(100\,\mathrm{AU}/r)^3\,(t/1\,\mathrm{Myr})$ m. Thus, 'boulders', i.e. bodies with sizes up to tens of metres, may form within 30 AU in a gas-clearance time-scale $\lesssim 1$ Myr, but probably much smaller 'icy dirtballs' farther out.

Steps to Making Comets – IV: Growth Without Gas

1. Dissipation of gas disc $\Longrightarrow$ further grain growth in absence of gas.
2. Two main channels: (1) standard 'planetesimal' picture and variants therein (widely accepted); (2) local gravitational instability picture (much less widely accepted — if at all!).
3. Consider 'standard picture':
   - $\Longrightarrow$ continued collisions and growth of 'boulders'/'snowballs' to bodies up to several tens of km in protoplanetary zone.
   - Produces comets with collisionally compressed structure on scale of 'boulders' (i.e. $\sim$10–100 m), and looser 'rubble-pile' structure on larger ($\gtrsim$ km) scales; $\Longrightarrow$ comets a collisionally evolved population.
   - Gravitational stirring by the largest bodies leads to continued growth, ultimately to make large planetesimals and planets.
4. Problems: time-scale to produce Uranus and Neptune too long; comets are planetary 'left-overs', formed in or close to outer planetary region; total cometary mass not much greater than that of the planets; role of EKB and how to form Oort cloud.

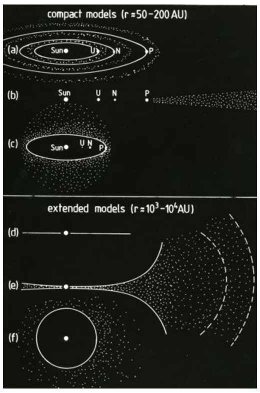
Steps to Making Comets – V: Gravitational Instability

1. Formation of dynamically cold, quiescent disc of icy dirtballs in outer solar system; random velocities decrease due to collisions.
2. $\Longrightarrow$ conditions for local gravitational instability: the 'icy dirtball' disc fragments into subdiscs with characteristic sizes λ_p depending only on Σ_d and r, i.e. $\lambda_p = 4\pi^2 G\Sigma_d/\Omega^2$, where $\Omega = (GM_\odot/r^3)^{1/2}$.
3. Subdiscs evolve like mini protoplanetary discs: to produce central objects (often multiple systems) with masses comparable to the mass m_p of the subdiscs, i.e. $m_p \simeq \pi\,(\lambda_p/8)^2\,\Sigma_s = \pi^5\Sigma_s^3 r^6/4M_\odot^2 \propto r^{3/2}$, i.e. $m_p \gtrsim 10^{18}$ kg for $r \gtrsim 50$ AU.
4. First-formed objects have masses comparable to observed outer solar-system objects; and — once formed — collisions become rare.
5. Predicts comets that are (1) mostly made in outer solar system by evolution of subdiscs to produce multiple central objects; (2) products of 'gentle' accumulation of 'boulders' or smaller 'snowball'-size components; and (3) largely collisionally unevolved.

Steps to Making Comets – VI: Summary

1. Comets produced by hierarchical accretion in outer planetary system; sizes range from a few km up to a few 100 km.
2. Cosmic Chemical Memory: interstellar and interplanetary dust aggregates contain ices that give clues to each of the distinct phases of grain growth in presence of gas, i.e. (1) interstellar gas and MC phases; (2) protostellar cloud and collapse phases; and (3) early disc evolution in presence of gas.
3. Evolution in absence of gas more uncertain; but 'boulders' and/or 'snowballs' must somehow grow into kilometre-size (and larger) comet nuclei.
   - In planetesimal picture, comets collisionally evolved; most formed in protoplanetary region and may have rubble-pile structure with more compact elements on scale of 'boulders' (10–100 m).
   - In gravitational instability picture, comets collisionally unevolved and of low-strength; most formed beyond planetary region and may have substructure on scale of 'snowballs' (metres or less).

New Theoretical Concepts: Dense Inner Cores



1. Compact Models:
   - (A) Interplanetary ‘rings’, left over from planet formation;
   - (B) a trans-Neptunian Edgeworth-Kuiper belt (EKB);
   - (C) a more spherical ‘Population II’ system.
2. Extended Models:
   - (D) a greatly extended, massive protoplanetary disc;
   - (E) a flared trans-Neptunian disc, merging smoothly into Oort cloud;
   - (F) a more spherical ‘Population II’ system.

Range of models leads to new ideas for origin of short-period comets.

New Discoveries: the Edgeworth-Kuiper Belt (EKB)

1. Discovery of Pluto (1930 February 18), announced March 13.
2. Consistent with earlier speculations (e.g. Lowell) about a ‘Planet X’ beyond Neptune; or that small objects might exist in the region beyond Neptune (e.g. Campbell 1916, Aitken 1926, Leuschner 1927, Leonard 1930).
3. Stimulates work by Edgeworth (1938, 1943); and later by Kuiper (1951), Whipple (1964), Fernández (1980), Duncan et al. (1988), Quinn et al. (1990) and others, focusing on JF short-period comets.
4. Searches by Kowal (1976–1985), Luu & Jewitt (1988), Levison & Duncan (1989), Tyson et al. (1992), eventually successful. Discovery of ‘QB1’, i.e. minor planet (15760) 1992 QB_1, the prototype ‘cubewano’ and the first ‘Kuiper Belt’ object.
5. Pluto = minor planet (134340) now among several trans-Neptunian objects (TNOs) classified as ‘dwarf planets’; Pluto: the ‘king’ of the Kuiper belt.

Pluto-Charon System and Edgeworth-Kuiper Belt

1. The classical ‘Kuiper Belt’: predicted by Irish scientist Kenneth Essex Edgeworth and others around middle of 20th century.
2. A region beyond Neptune, comprising a vast swarm, or belt, of icy planetesimals in low-inclination orbits.
3. $\approx 10^5$ objects with diameters greater than 100 km. Many more ($\approx 10^9$), it is believed, of ‘ordinary comet’ size.

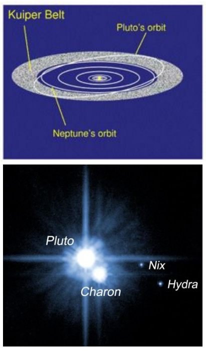


Image credits: Johns Hopkins University; NASA/ESA/HST

Artist’s Impression of Some Large TNOs/Dwarf Planets

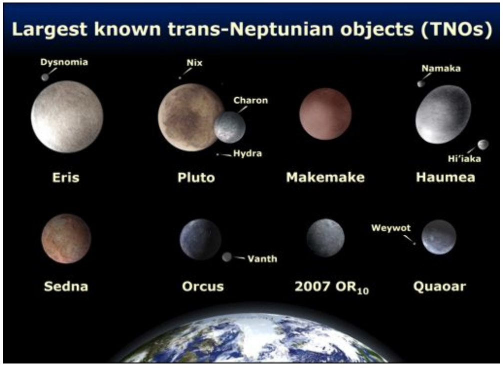


Image Credit: Wikipedia; based on 2006 press release by NASA/ESA/HST.

Evolution of Oort Cloud: Contrasting Views

1. 1950 Model: quasi-steady state; comets in long-term ‘deep freeze’ of Oort cloud for age of solar system; stars dominate the evolution; no dense inner core.
2. Modern view:
   - On short timescales: $t \lesssim 10$ Myr: Changes in perihelion distance dominated by Galactic tide and passing stars; leads to a quasi-steady long-period comet flux; changes in orbital energy dominated by stellar perturbations.
   - On medium timescales: $10 \lesssim t \lesssim 500$ Myr: periodic new-comet flux due to Sun’s orbit about Galactic plan; rare, close stellar passages more important for randomizing orbits; changes in orbital energy still dominated by stars.
   - On long timescales: $500 \lesssim t \lesssim 4000$ Myr: rare, close molecular cloud encounters disrupt outer cloud, dominating changes in orbital energy beyond c.10,000 AU; rare, close stellar encounters stir up inner core.

   Together, these major upheavals replenish the transition zone between inner and outer Oort cloud and stir up orbits in Dense Inner Core.

Galactic Impact: Time-Variable Cometary Flux

1. Galactic tide dominates quasi-steady new-comet flux from Oort cloud.
2. Comet flux roughly proportional to mass-density, $\rho_S(t)$ at Sun’s location in Galaxy (see Figure, after J. Matese et al. 1995).
3. Δq per revolution depends on q, a, and Galactic latitude of perihelion, b, i.e.

Δq per revolution $=$ $(10\pi^2\sqrt{2}\,\rho/M_\odot)\sin(2b)\,q^{1/2}a^{7/2}$

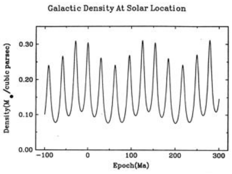
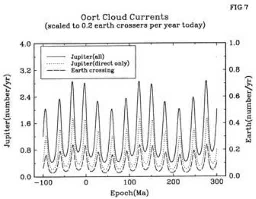
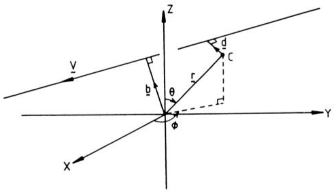


FIG 7



Survival Problem: Physics of External Perturbations

1. Consider a perturber of mass M passing Sun with velocity $\mathbf{V}$ and impact parameter $\mathbf{b}$ with respect to Sun and $\mathbf{d}$ with respect to a comet at heliocentric distance $\mathbf{r}$.



2. Then the relative velocity change of the comet with respect to the Sun is the difference of the two impulses, i.e.

$$\mathbf{\Delta v} = \frac{2GM}{dV}\hat{\mathbf{d}} - \frac{2GM}{bV}\hat{\mathbf{b}} = \frac{2GM}{bV}\left\{\left(\frac{b^2}{d^2} - 1\right)\hat{\mathbf{b}} - \frac{rb}{d^2}\left[\hat{\mathbf{r}} - (\hat{\mathbf{r}}.\hat{\mathbf{V}})\hat{\mathbf{V}}\right]\right\}$$

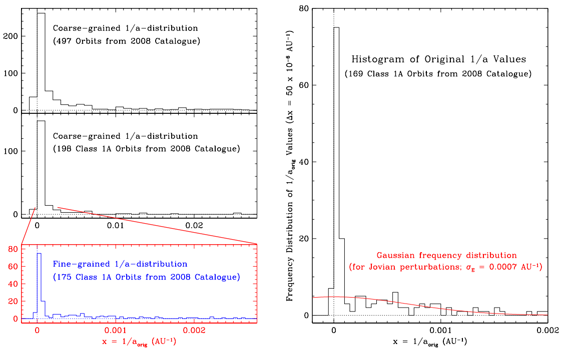
Summary of Survival Problem: Oort Cloud Evolution

1. Two main types of external perturber: stars and molecular clouds.
   - Galactic tide also drives comets into inner solar system, but has little direct effect on Oort cloud’s disruption.
2. Stars pass through and beyond the Oort cloud, causing gradual unbinding of cometary orbits; the ‘stellar’ half-life is $t_{1/2,*} \simeq 2 \times \left(2 \times 10^4\,\mathrm{AU}/a\right)$ Gyr.
3. Molecular clouds pass beyond the Oort cloud, but are much more massive than stars; the ‘molecular cloud’ half-life is $t_{1/2,c} \simeq 2 \times \left(2 \times 10^4\,\mathrm{AU}/a\right)^3$ Gyr

$\Longrightarrow$ ‘standard Oort cloud dynamically unstable beyond $a \simeq 2 \times 10^4$ AU, over the age of the solar system (4.5 Gyr)

The Oort cloud is a leaky reservoir which must be replenished from within, possibly the trans-Neptunian region or a Dense Inner Core.

Fading Problem: Recall $1/a$-distribution



Halley-Type Comet (HTC) Capture Probability

Inclination-averaged mean capture probability from near-parabolic orbits to a HTC orbit:

- Decreases sharply with increasing q.
- Non-zero out to $q \simeq 15\,\text{AU}$.
- Averages $\sim$0.01 for $q \lesssim 5\,\text{AU}$.

The new-comet flux ($\sim$1 per year) and mean dynamical lifetime as a HTC ($\sim$0.3 Myr), and the capture probability, p_c determines the predicted number of HTCs.

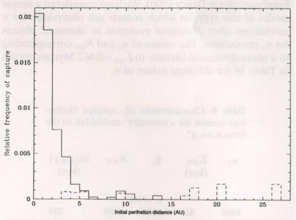


Figure 4. The relative frequency of capture from an isotropic near-parabolic source into Halley-type orbits (solid line) and Jupiter-family orbits (dashed line). Note that whereas Halley-type comets primarily originate from orbits of small perihelion distance, Jupiter-family comets primarily come from initial orbits in the outer Solar system.

Fading Problem: Where Are the 'Dead' Comets?

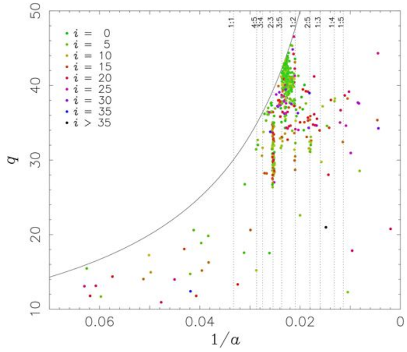
1. Observed new-comet flux: Approximately 1 comet per year brighter than $H_{10} = 7$ (corresponds to diameter $d \gtrsim 5\,\text{km}$) with $q < 5\,\text{AU}$, i.e. with perihelion distance within Jupiter's orbit.
2. Capture probability to 'Halley-type comet' (HTC), i.e. capture probability to $P \lesssim 200\,\text{yr}$: $p_c \simeq 0.01$ per new comet; the rest get ejected.
3. Mean dynamical lifetime as a Halley-type comet: $t_{\text{dyn}} \simeq 3 \times 10^5\,\text{yr}$.
4. $\Longrightarrow$ steady-state number of HTCs, N_{HTC}, given by $N_{\text{HTC}} \simeq 1 \times 0.01 \times 300,000 \approx 3000$.
5. 30–100 times more than observed: where are the dead comets?!
    - Perhaps they are 'dark' HT asteroids; 'boulders'; or 'dust'?
    - In any case, comets must have short lifetimes in visible region.

Origin of Jupiter-Family Comets? — Not Kuiper Belt!

1. Jupiter-family (JF) short-period comets (SPCs) mostly have low inclinations.
2. Suggests a source in a flattened, low-eccentricity, disc-like distribution (Duncan, Quinn, & Tremaine, 1988): the 'classical' Kuiper belt.
    - Comets must have perihelia near Neptune, i.e. $q \approx 30\,\text{AU}$, in order to be efficiently captured and 'handed down' to the Jupiter family.
    - Simulations require at least 4×10^9 such comets in the comet belt, if this is the dominant source of JFCs.
    - The dynamical lifetime of JFCs is $\approx 3 \times 10^5\,\text{yr}$; their active lifetime is much shorter, i.e. $\approx 1.2 \times 10^4\,\text{yr}$ (otherwise inclinations increase).
3. Two main problems: (1) the required source orbits are not observed; and (2) evolution of an initial distribution of low-inclination Neptune-crossing orbits inevitably produces a 'Scattered Disc' containing a similar number of comets in much more eccentric low-inclination orbits. These more readily captured into JFC orbits.

$\Longrightarrow$ JFCs primarily not from Kuiper belt, but from Scattered Disc.

Distribution of Observed High-Accuracy TNOs



High-accuracy TNO orbits showing mean-motion resonances with Neptune and lack of non-resonant objects with q near Neptune and $a \lesssim 50\,\text{AU}$. Image credit: David Asher.

New Number Problems: Scattered Disc and Oort Cloud

1. Simulations require number of comets in Scattered Disc to be $\lesssim 10^9$.

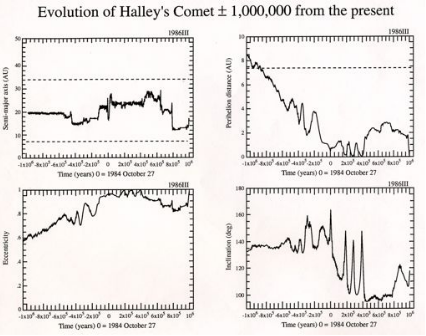
    But best observational estimates of number in Scattered Disc are ≈ 1–2×10^8, albeit with large uncertainty.
2. Simulations require number of comets in Oort Cloud to be approximately ten times the number in Scattered Disc, i.e. $\lesssim 10^{10}$.

    But best observational estimates of number in Oort Cloud $\gtrsim 2 \times 10^{11}$.
3. Two recent 'solutions':
    - Observed JFCs not in steady state; or large Scattered Disc Objects tidally break up into many fragments during dynamical evolution towards JFCs (Volk & Malhotra 2008).
    - Observed Oort cloud not primordial to solar system, but instead comprises largely captured comets, ejected from the Scattered Discs of other stars making up our Sun's parent star cluster (Levison, Duncan, Brasser & Kaufmann 2010).

New Discoveries: Complex Dynamical Evolution – I

e.g. 1P/Halley:

1. Resonances: mean-motion and secular.
2. Kozai Cycles: Correlated large changes of eccentricity and inclination.
3. Sungrazing: a ubiquitous cometary end-state.


Evolution of Halley's Comet ± 1,000,000 from the present

New Discoveries: Complex Dynamical Evolution – II

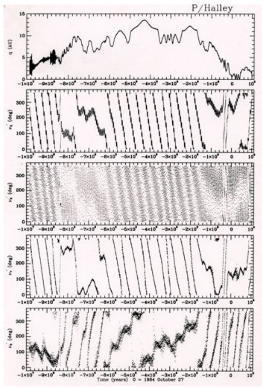
e.g. 1P/Halley Secular Resonances:

1. Note enormous secular evolution of perihelion distance.

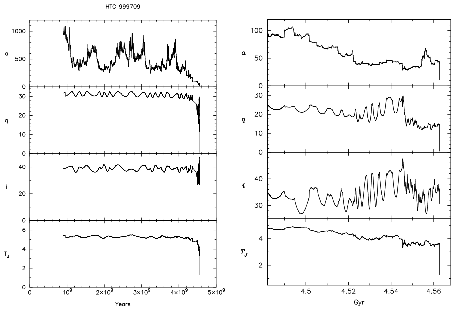
    Associated with critical argument $\nu_p = (\Omega - \omega) - (\Omega_p + \omega_p)$

    When this $\sim$constant, line of apses of comet's orbit locks on to the rate of precession of one of the giant planets (J, S, U, N).

    Figure shows effects on q of such resonances with Jupiter (2nd panel) and Neptune (5th panel). $-10 < t < 1\,\text{Myr}$.
2. This kind of evolution totally unexpected: quite different from pure 'random walk'.



Halley-Type Comets From Inner Oort Cloud



Example of Halley-type comet from inner Oort cloud, involving gradual dynamical transfer from outer solar system ($a > 10^3$ AU and initial q near Neptune) through weak perturbations. ≈10% of HTCs originate this way. Image credit: Emel'yanenko et al. 2007, MNRAS, 381, 779–789.

Persistent Puzzles – II

1. How are comets formed? And where?

   In or beyond the protoplanetary region; in the Sun's parent molecular cloud; or interstellar space?

   - Standard View: by coagulation and/or subsequent gravitational instability of small bodies (ice-covered dust/boulders) in and/or just beyond protoplanetary disc.

2. What are the proximate source(s) of observed comets?

   - Standard View: Principally Oort Cloud for long-period and Halley-type SPCs; the Scattered Disc for Centaurs and Jupiter-family SPCs.
   - Another View: Principally Oort Cloud for long-period, Halley-type SPCs, 'Centaurs' and ∼50% the JFCs; the observed near-Neptune high-eccentricity (NNHE) region for the remaining ∼50% JFCs.

   On this view, 'Centaurs' have $5 < q < 28$ AU and $a < 1000$ AU, and any inclination, i; NNHE objects have $28 < q < 35.5$ AU and $60 < a < 1000$ AU; also any i.

Persistent Puzzles – III

1. Comets very diverse. But are they essentially the same object, or are there two or more subtypes, e.g. correlating with orbital period or class?
2. Standard View: At least two subtypes, principally Centaurs and JFCs, both from the Scattered Disc (mostly from the near-Neptune part of the protoplanetary disc); and LPCs and HTCs from the Oort Cloud (mostly from the Jupiter-Saturn-Uranus region).
   - JFCs have long active lifetimes in the visible region ($q< \lesssim 2.5$ AU), greater than ≈ 10^3 revolutions; LPCs and HTCs have short active lifetimes, less than ≈200 revolutions.
3. Alternative View: All comets essentially the same: formed in outer regions of heterogeneous protoplanetary disc and subsequently ejected to produce Oort cloud and more flattened dense inner core.
   - All comets fragile; short active lifetimes, less than 200 revolutions, in visible region. No distinction between HTCs and most JFCs.

Further Questions

1. What is structure and evolution of the 'observed' Oort cloud. How was it formed; and does it contain a massive dense inner core?
2. What is the cometary mass function and the average cometary mass? How many comets are there, and what is their total mass?
   - Consistent with a standard 'low mass' protoplanetary disc?
3. What is role played by newly discovered outer solar system bodies: Centaurs, Edgeworth-Kuiper belt objects, trans-Neptunian objects etc?
4. Are comets fragile or strong? What are their end-states; and what is the impact of cometary debris on the Earth, other planets, and Sun?

Conclusions — General Points

1. 'Comets' can sometimes be the most prominent objects in sky; their study goes back thousands of years.
2. Comets touch on many areas of astronomy, not least solar-system science; they have had a significant impact on the Earth and on the development of scientific ideas.
3. Earth an 'open' system, in touch with its near-space celestial environment: a paradigm shift as significant as Copernicanism.
4. Solar system 'very leaky'; interesting implications for the dust, small bodies and planets in molecular clouds and the interstellar medium; what about comet clouds around other stars?
5. Modern picture of comets; a balance between the historical catastrophist and Newtonian uniformitarian views; comets as potential destroyers of life and as objects that bring the necessities of life (e.g. water, organics, perhaps seeds of life itself) to Earth.

Conclusions — Some Problems to be Resolved

1. Ancient history suggests 'the sky' may have been significantly different in proto-historic times (e.g. more 'active', more interplanetary debris, brighter zodiacal light etc.); how can that be?
2. Cometary masses range up to the size of dwarf planets; what are the effects of occasional 'giants' on Earth (and Sun)? What is the average mass of a comet?
3. Are all comets essentially the same; or are there two or more different classes, e.g. depending on origin and/or dynamical characteristics?
4. Total mass of Oort cloud may be very large (≈ $10^2\,M_\oplus$ pc^{-3}); implies potentially serious difficulties for 'standard' primordial solar system picture?
5. 'Fading problem' still not understood, but effectively determines the predicted $1/a$-distribution; what happens to the cometary debris?
6. Meteoroid streams initially very fine-grained; leads to strong time-dependence in accretion of dust and small bodies on Earth.

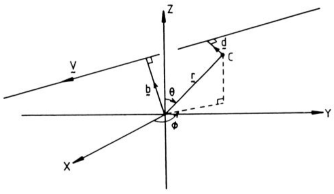
Acknowledgements

Astronomy at Armagh Observatory is funded by the Northern Ireland Department of Culture, Arts and Leisure

Department of
Culture, Arts
and Leisure
www.dcalni.gov.uk

Survival Problem: Physics of External Perturbations

1. Consider a perturber of mass M passing Sun with velocity $\mathbf{V}$ and impact parameter $\mathbf{b}$ with respect to Sun and $\mathbf{d}$ with respect to a comet at heliocentric distance $\mathbf{r}$.



2. Then the relative velocity change of the comet with respect to the Sun is the difference of the two impulses, i.e.

$$\mathbf{\Delta v} = \frac{2GM}{dV}\hat{\mathbf{d}} - \frac{2GM}{bV}\hat{\mathbf{b}} = \frac{2GM}{bV}\left\{\left(\frac{b^2}{d^2} - 1\right)\hat{\mathbf{b}} - \frac{rb}{d^2}\left[\hat{\mathbf{r}} - (\hat{\mathbf{r}}.\hat{\mathbf{V}})\hat{\mathbf{V}}\right]\right\}$$

Heuristic Results

1. Mean relative velocity change in a single encounter is approximately:

$$\Delta v = \frac{2GM}{bV}\begin{cases} \sqrt{2} & b < \sqrt{7/12}\,a \\ \sqrt{7/6}\,a/b & b > \sqrt{7/12}\,a \end{cases}$$

2. On short timescales (e.g. $t \lesssim 30\,\mathrm{Myr}$), the closest stellar encounter expected during a given time interval t has impact parameter $b_{\mathrm{min}} \simeq (2\pi nVt)^{-1/2}$, where n is the number density of perturbers. For stars this usually implies $b \gtrsim a$, which leads to

$$\Delta v_{\mathrm{max}} \simeq 4\pi(7/6)^{1/2}\, G\rho a t$$

where $\rho = nM$ ($\approx 0.05\,M_{\odot}\ \mathrm{pc}^{-3}$ for stars) is the mass density of perturbers.

3. This leads to $\Delta v_{\mathrm{max}} \simeq 4.3\,(a/3 \times 10^4\,\mathrm{AU})(t/10\,\mathrm{Myr})\ \mathrm{m\ s}^{-1}$.

4. Finally, setting $t = P(a) \simeq 5.2\,(a/3 \times 10^4\,\mathrm{AU})^{3/2}\,\mathrm{Myr}$, the maximum change in perihelion distance during a single revolution can be shown to be of the order of

$$\Delta q \approx 5\,(a/3 \times 10^4\,\mathrm{AU})^{7/2}(q/1\,\mathrm{AU})^{1/2} \quad \mathrm{AU}$$

Mean Energy Transfer Rate

1. Change in orbital energy in a single encounter: let $\mathbf{\Delta v}$ be the relative velocity change of the comet with respect to the Sun, and let $\mathbf{v_0}$ be its orbital velocity, then

$$\Delta E = \mathbf{v_0}.\mathbf{\Delta v} + \frac{1}{2}\,(\mathbf{\Delta v})^2$$

Cometary orbital energies thus diffuse and systematically increase (i.e. become less tightly bound) owing to external perturbations.

2. Approximate result for point-mass perturbers: define $a_c = \sqrt{12/7}b_{\mathrm{min}}$, where $b_{\mathrm{min}} \simeq (2\pi nVt)^{-1/2}$ is the most probable minimum impact parameter for the perturbers of number density n, then the mean energy transfer rate can be shown to be approximately

$$\dot{\varepsilon}(t) = \frac{4\pi G^2 M^2 n}{V}\begin{cases} (a/a_c)^2 & a < a_c \\ 2\ln(a/a_c) + 1 & a > a_c \end{cases}$$

Time-Scales For Survival

1. For stars and $t \simeq 4.5\,\mathrm{Gyr}$, we have $a \gg a_c$. This implies $\dot{\varepsilon}_* \approx$ const..
2. For molecular clouds and $t \simeq 4.5\,\mathrm{Gyr}$, we have $a \ll a_c$, i.e. $\dot{\varepsilon}_c \propto a^2$.
3. The net result is:

$$\dot{\varepsilon} = \dot{\varepsilon}_* + \dot{\varepsilon}_c \simeq A_* + A_c a^2$$

where for typical parameters $A_* \simeq 10^{-13}\ \mathrm{m^2 s^{-3}}$ and $A_c \approx 10^{-44}\ \mathrm{s^{-3}}$.

4. Solving the energy evolution equation for each type of perturber leads to the half-life due to stellar and molecular cloud perturbations, i.e.

$$t_{1/2,*} = \frac{1}{4.732}\frac{GM_{\odot}}{A_* a} \simeq 2 \times 10^9 \left(\frac{2 \times 10^4\,\mathrm{AU}}{a}\right)\ \mathrm{yr}$$

and

$$t_{1/2,c} = \frac{1}{8.190}\frac{GM_{\odot}}{A_c a^3} \simeq 2 \times 10^9 \left(\frac{2 \times 10^4\,\mathrm{AU}}{a}\right)^3\ \mathrm{yr}$$

5. Thus, due to both clouds and stars, the majority of comets with initial $a \gtrsim 2 \times 10^4\,\mathrm{AU}$ will be lost. This is the Oort cloud survival problem.

Summary of Survival Problem: Oort Cloud Evolution

1. Two main types of external perturber: stars and molecular clouds.
    - Galactic tide also drives comets into inner solar system, but has little direct effect on Oort cloud's disruption.
2. Stars pass through and beyond the Oort cloud, causing gradual unbinding of cometary orbits; the 'stellar' half-life is $t_{1/2,*} \simeq 2 \times (2 \times 10^4\,\mathrm{AU}/a)$ Gyr.
3. Molecular clouds pass beyond the Oort cloud, but are much more massive than stars; the 'molecular cloud' half-life is $t_{1/2,c} \simeq 2 \times (2 \times 10^4\,\mathrm{AU}/a)^3$ Gyr

$\Longrightarrow$ 'standard Oort cloud dynamically unstable beyond $a \simeq 2 \times 10^4$ AU, over the age of the solar system (4.5 Gyr)

The Oort cloud is a leaky reservoir which must be replenished from within, possibly the trans-Neptunian region or a Dense Inner Core.