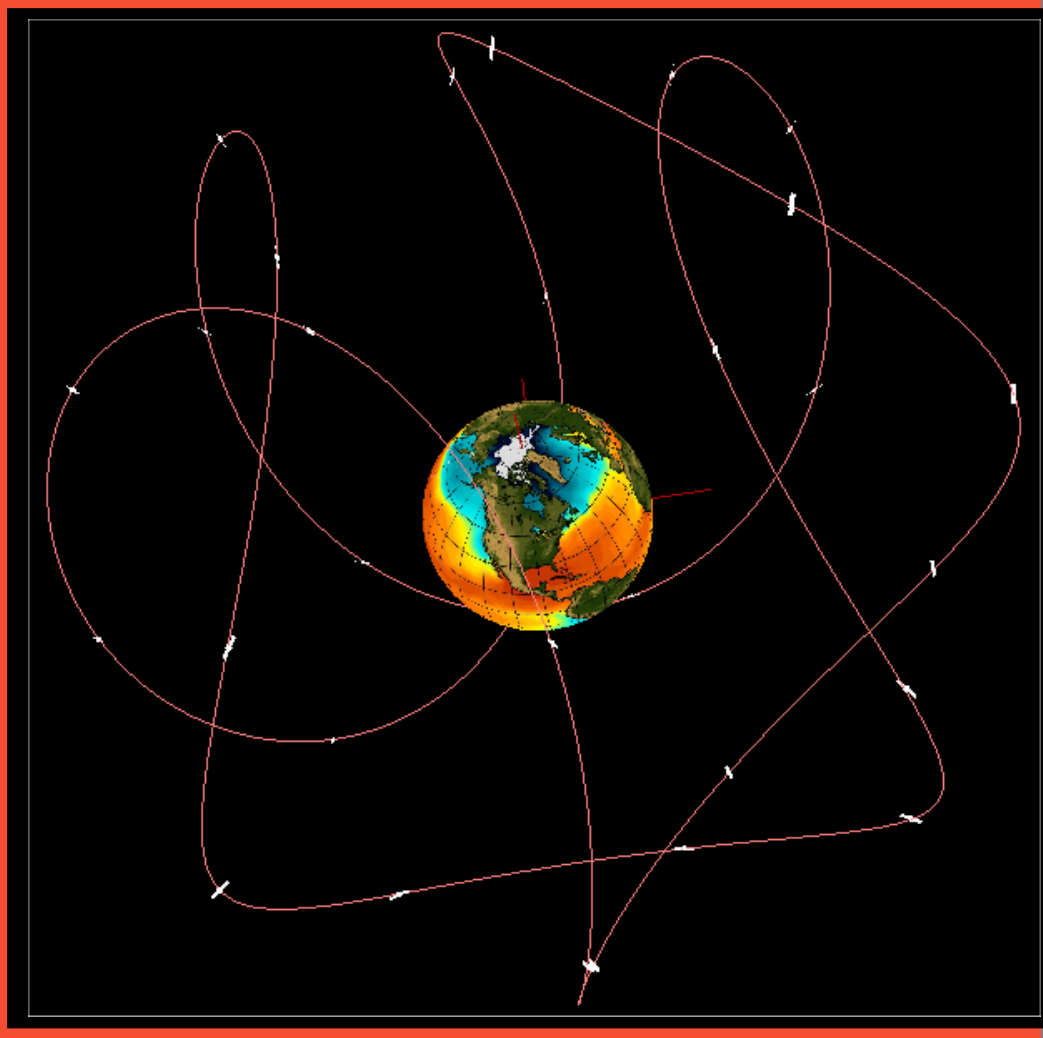
ACT-PUB-AF0502

Acta Futura - Issue 2

Special Issue on Innovative System Concepts

Editor: Dr. Dario Izzo

Acta Futura: From ESA's Advanced Concepts Team



Advanced Concepts Team

http://www.esa.int/act

European Space Agency
Agence spatiale européenne

Contents

Foreword

New ideas are often slow to penetrate mainstream space science. They usually require a lengthy gestation period to become accepted as practical tools in the design of real missions. For example, discussions in scientific literature on subjects such as formation flying or electric propulsion took place years before their implementation as methods to augment capabilities and lower costs. With a critical eye, it must therefore be possible to sift through the vast body of emerging research to find the ideas that will one day take their place in space mission design. The role of ESA's Advanced Concepts Team (ACT) is to identify and nurture such ideas, to identify new useful concepts and help them grow into pervasive tendrils in our community.

It is with this purpose in mind that the ACT organised the 1st Workshop on Innovative System Concepts, on the 21st of February 2006 at the European Space and Technology Center in Noordwijk, The Netherlands. Invited speakers discussed ways of using spindly tethers that pull electrical power out of magnetic fields to explore the moons and planets of the outer Solar System where little is available. Some of the missions considered used advanced propulsion options to send spacecraft to deflect dangerous asteroids or to go beyond the solar system boundaries. Some speakers showed how the use of a swarm of tiny satellites in formation may synthesise large structures such as telescopes and sails. Others proposed to code primitive goal-oriented instincts into the control system of each satellite of a group to guide it to complete simple tasks while remaining unaware of contributing to a more complex undertaking. The workshop audience could also learn how advances in constellation design theory allow now to consider constellations of satellites that behave like rigid objects in space, seemingly mocking Keplers Laws.

Some aspects of the ideas presented during the workshop are, or have been, also subject of studies under ESA's Ariadna programme. The main purpose of Ariadna is to enhance cooperation and facilitate research partnerships between ESA and investigators in universities and research groups.

While envisaged to give neither a definitive nor comprehensive blueprint of the future, the workshop has surely invigorated the debate on new ideas for space engineering. This second issue of *Acta Futura* offers a compilation of papers written by the speakers of the Workshop on Innovative System Concepts. Each one of the papers contains an original contribution to some advanced research area discussed at the workshop and agreed to have potentials to be, one day, part of the common background of space engineers. By pulling the right threads, the convoluted web of current research may be unraveled to help build a path toward ESA's objectives.

This issue of Acta Futura is published in the form of an ESA special publication and has to be referenced as such.

More publications and studies are available from the ACT website.

Dario Izzo, Nicholas Lan
Advanced Concepts Team

http://www.esa.int/act
act@esa.int

Electrodynamic Tethers for Exploration of Jupiter and its Icy Moons

Juan R. Sanmartin

ETSI Aeronáuticos, Universidad Politécnica de Madrid, Pza. C. Cisneros 3, 28040 Madrid, Spain, jrs@faia.upm.es

Abstract

Use of electrodynamic bare tethers in exploring the Jovian system by tapping its rotational energy for power and propulsion is studied. The position of perijove and apojove in elliptical orbits, relative to the synchronous orbit at 2.24 times *Jupiter*'s radius, is exploited to conveniently make the induced Lorentz force to be drag or thrust, while generating power, and navigating the system. Capture and evolution to a low elliptical orbit near *Jupiter*, and capture into low circular orbits at moons Io and Europa are discussed.

1. Introduction

A full study of the giant, complex Jovian system is a central goal in space science [1]. There exists a pressing need for a spacecraft (S/C) to reach into a low orbit around moon *Europa*, as well as around moon Io and *Jupiter* itself. Within such a scope, the successful Galileo mission was a handcuffed mission. The need of gravity assist manoeuvres (GAMs) to reach *Jupiter* resulted in quite restricted launch windows and a protracted trip. Just the capture operation required too much chemical propellant, reducing scientific payload to a few percent in mass. The power source used, Radioisotope Thermal Generators (RTGs), is too weak. Capabilities for orbit manoeuvring after capture and for data transmission were very low [2].

In 1999 the US National Research Council made full scientific planning for a mission to *Europa*, which would still use RTG's for power, and use gravity assists for a Jovian moon tour supposedly acquiring a low *Europa* orbit in a few months [3]; escalating costs, however, made NASA cancel the mission in 2002. At about that time, NASA embarked in *Project Prometheus* on the use of nuclear reactors for both power, and powering electrical thrusters (NEP). Original planning for a *Jovian Icy Moons Orbiter* (an "unfriendly", 20 ton system), gave later place to *Juno* (polar Jovian) and *Neptune-Triton* missions.

ESA in turn has made plans about a *Jovian Minisat Explorer*, which would keep some *Galileo* features (GAMS-determined trip from *Earth*, chemical-rocket capture) but move back from RTG's to solar arrays. ESA considers developing *Low Intensity Low Temperature* cells with solar concentrators. Should that program fail, ESA would revert to RTG's, which are socially problematic however. Also, Pu-238 oxide is scarce and expensive (10^6€/ kg) and ESA must face ITAR restrictions on US RTG technology. On reaching *Jupiter*, the *JME* would split into a *Jovian Relay Satellite* and a *Jovian Europan Orbiter* due to acquire a low *Europa* orbit through an extended series of GAMS in 550 days [4].

The approach here discussed would involve neither RTG's, nor nuclear reactors or solar arrays. It would use an electrodynamic (ED) tether system, accounting for a moderate fraction of S/C mass, to tap *Jupiter*'s rotational energy for **both** power and propulsion [5, 6]. It should result in a direct trip from *Earth* and higher data-handling and scientific payload capabilities, and it would allow for a fast manoeuvring, 'free-lunch' tour (using GAMs and chemical propulsion very sparingly). Since tether performance is dependent on ambient conditions (magnetic field B and plasma density N_e), the critical phase is S/C capture. Ambient model uncertainties would suggest launching two light S/C, one designed for nominal conditions, the other with greater design margins; as a bonus, they would make possible determining spatial structure in the extensive Jovian magnetosphere.

2. Power Generation, Drag and Thrust at an ED-Tether

Consider a simple planet/light-satellite system. Both planetary spin and orbital motion contribute to mechanical energy,

$$\epsilon_{mech} = \frac{1}{2} I_{pl} \omega_{pl}^2 - \frac{\mu_{pl} M_{sat}}{2a} \tag{1}$$

and to angular momentum

$$H = I_{pl} \omega_{pl} + \frac{\mu_{pl} M_{sat}}{2a \Omega_{orb}} \tag{2}$$

With H = const $\equiv H_0$, Kepler's law $a^3 \Omega_{orb}^2 = \mu_{pl}$ determines $\epsilon_{mech}(a; H_0)$,

$$\frac{2a_* \epsilon_{mech}}{\mu_{pl} M_{sat}} = \left(\frac{H_0}{I_{pl} \sqrt{\mu_{pl}/a_*^3}} - \sqrt{\frac{a}{a_*}} \right)^2 - \frac{a_*}{a} \tag{3}$$

where a_*^2 is $\frac{I_{pl}}{M_{sat}}$ and where we assumed a circular equatorial orbit with $\bar{\mathbf{H}}_\mathbf{0} \cdot \bar{\mathbf{\Omega}}_\mathbf{orb} > 0$.
If $3^{3/4} H_0 > 4 I_{pl} (\mu_{pl}/a_*^3)^{1/2}$, the graph $\epsilon_{mech}(a)$ presents a maximum, and a minimum farther out, both extrema corresponding to rigid-body motion, $\Omega_{orb} = \omega_{pl}$. The maximum is always unstable, any kinetic mechanism for dissipation (specifically, tidal forces) would drive the satellite away from rigid-body motion at a(max), on either side of it. For the comparatively extremely light artificial satellites, the maximum lies at the synchronous or stationary radius $a_{s,pl} \equiv (\mu_{pl}/\omega_{pl}^2)^{1/3}$ where a satellite corotates with the planet (the energy minimum lying beyond Universe limits even for multi-ton space stations). If one would have a corotating atmosphere beyond a_s, satellites at $a > a_s$ would be pushed by faster-moving air to higher (though slower) orbits, air thus exerting thrust rather than drag.
In planets that have both magnetic field and ionosphere/magnetosphere, an orbiting conductive tether provides an alternative dissipative mechanism. Consider the nonrelativistic equation for transformation of electric field,

$$\begin{aligned} \bar{E}(tether frame) - \bar{E}(plasma frame) \\ = \bar{E}_m \equiv (\bar{\nu}_{orb} - \bar{\nu}_{pl}) \wedge \bar{B} \end{aligned} \tag{4}$$

In the highly conductive plasma away from the tether, the electric field will be negligible in the frame moving with the corotating plasma, yielding, in the tether frame, $\bar{E}(outside) = \bar{E}_m$. This outside field will drive a current inside the tether, $\bar{I} \propto \bar{E}$ (inside, in tether frame), with $\bar{I} \cdot \bar{E}_m > 0$. Using Eq. 4 for E_m and the Lorentz force $L\bar{I} \wedge \bar{B}$, the net mechanical power in the tether-plasma interaction becomes

$$L\bar{I} \wedge \bar{B} \cdot (\bar{\nu}_{orb} - \bar{\nu}_{pl}) = -L\bar{I} \cdot \bar{E}_m < 0 \tag{5}$$

that lost power appearing in the tether electrical circuit. Clearly, $L\bar{I} \wedge \bar{B} \cdot \bar{\nu}_{orb}$ will be positive, corresponding to thrust acting on the tethered S/C, if $\bar{\nu}_{orb}$ is opposite $\bar{\nu}_{orb} - \bar{\nu}_{pl}$; this recovers the $a > a_s$ condition [6]. Figure 1 illustrates this condition for the tether-plasma interaction.
The basic requirement for quasisteady ED-tether operation is establishing effective contact, both anodic and cathodic, with the ambient plasma. Electron ejection is not an issue. Hollow cathodes are presently reaching ratios of current to expellant mass-flow rate as large as $10^2 A/mgs^{-1}$ (which is about the charge/mass ratio of protons). This results in fully negligible expellant consumption at a hollow cathode (HC): For $B \sim 1$ Gauss and a tether length L of tens of kilometres, the ratio of the Lorentz force to the expellant mass-flow-rate is well over 10,000 km/s, which is several orders of magnitude larger than the jet velocity of electrical thrusters.
As regards the problem of anodic contact with a highly rarefied plasma, it was solved in 1992 when it was proposed that, instead of using a big end collector, the tether be left bare of insulation, allowing it to collect electrons - as a cylindrical Langmuir probe in the orbital motion limited (OML) regime - over the segment coming out polarized positive [7]. A length-averaged tether current, $I_{a\nu}$, should now figure in the Lorentz force. The collecting area of a thin bare tether can be large because that segment may be tens of kilometres long. Collection can be efficient if the cross section dimension is smaller than both electron Debye length and gyroradius [8]. The cylindrical geometry allows a final bonus; a thin tape can collect the same current as a round wire of equal cross-section perimeter [9] and will be much lighter. The optimal tether is thus characterized by three quite disparate dimensions, $L \gg w$ (width) $\gg h$ (thickness).

3. The Jupiter Free-Lunch Tour

The Jovian system is a particularly appropriate place for using an ED-tether. The stationary orbit for a planet is readily shown to satisfy the relation

$$a_s / R_{pl} \propto (\rho_{pl} / \Omega_{pl}^2)^{1/3} \tag{6}$$

Jupiter has both low mean density ρ_J and rapid rotation; as a result the stationary orbit lies at $a_{sJ} \approx 2.24 R_J$, which is one third of the relative distance for *Earth*. Further, the surface magnetic field is ten times greater at *Jupiter* than at *Earth*, magnetic pressure and tension thus being 10^2 times greater in *Jupiter*. A plasmasphere reaches to about $3.8 R_J$, well beyond a_{sJ}. [10]
In addition, moon *Io* is both at a 1:2 *Laplace* resonance with *Europa*, and ten times relatively closer to its planet than the Moon is to *Earth* ($a_{Io} \sim 5.89 R_J$). This leads to tidal deformations inside *Io* that produce extreme tectonics and volcanism. Neutral gas continuosly ejected by *Io* is ionized and accelerated by the fast-flowing Jovian magnetosphere, and made to corotate as a giant plasma

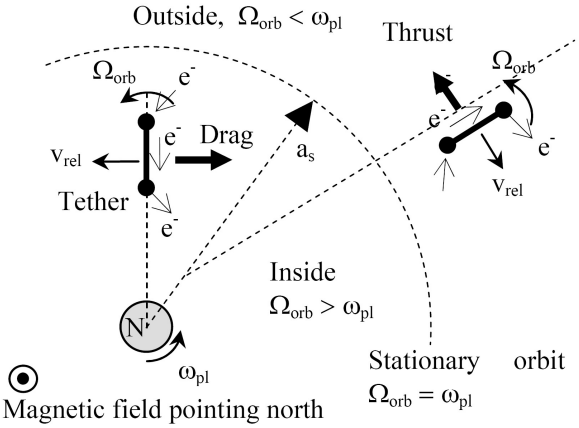


Figure 1: Tether operation inside/outside 'drag sphere'

torus, which is denser than the plasmasphere and reaches from about the plasmasphere to *Europa*, orbiting at $a_{Eu} \approx 9.38R_J$.

Tether drag/thrust will only be exerted within plasmasphere or torus; the tether current can be (nearly) shut off at convenience by switching off the HC or plugging a large resistance in the tether circuit. The proposed Jovian tour will exploit the positions of perifocus in the orbit coming from *Earth* and of perijove and apojove after capture, relative to the 'drag sphere', to exert either drag or thrust; notice that this sphere only roughly indicates whether drag or thrust applies in case of noncircular orbits.

The Jovian tour starts with a S/C approaching *Jupiter* at the relative velocity $\nu_\infty \sim 5.7km/s$ of a minimum energy transfer. Assuming the perifocus is at $r_p \sim 1.5R_J$ the hyperbolic excentricity is very small, $e-1 = \nu_\infty^2 r_p/\mu_J \approx 0.027$. After capture, closed orbits evolve under repeated Lorentz force, as schematically shown in Fig. 2.

Spacecraft capture requires drag to make a minimum work, this condition being roughly written as

$$\alpha \times LI_{a\nu}B \times \pi r_p = (1+\beta) \times \frac{1}{2} M_{S/C} \nu_\infty^2 \qquad (7)$$

If ohmic and HC contact impedances are neglected, the tether will be biased positive throughout its length and the averaged tether current is 2/5 of the OML current [7] at uniform bias $E_m L$,

$$I_{a\nu} = \frac{2}{5} \times \frac{2wL}{\pi} eN_e \sqrt{\frac{2eE_m L}{m_e}} \qquad (8)$$

Introducing the tape mass $m_t = \rho_t Lwh$, yields a mass ratio condition,

$$\frac{M_{SC}}{m_t} = \frac{8\alpha/5}{1+\beta} \times \frac{m_e N_e r_p \sqrt{2eE_m L/m_e} \times LeB/m_e}{\rho_t h \times \nu_\infty^2}$$

We take values $N_e \approx 10^3$ cm^{-3}, $B \approx 1.6$ gauss, $E_m \approx 4.8$ V/m at the perifocus r_p from the *Divine-Garrett* Jovian model [11], and introduce a factor $\alpha = 0.5$ to roughly account for variations along a drag path $\sim \pi r_p$. Capture requires a positive β; the greater is β the lower is the eccentricity of the capture orbit. Using $\beta = 3/4$ and an *Al* tape of thickness $h = 0.05$ mm and length $L = 50$ km, yields

$$M_{SC}/m_t \approx 4.15 \qquad (9)$$

For β barely positive, the mass ratio would be about 7.25. Note that the ratio M_{SC}/m_t is independent of tape width w and increases with the ratio $L^{3/2}/h$. There is a limit, however, to the possible gain in mass ratio because of the OML-to-short circuit current ratio

$$\frac{I(OML)}{I(shortcircuit)} \propto \frac{N_e\sqrt{E_m} \times wL^{3/2}}{\sigma_{cond} E_m wh} \qquad (10)$$

is also proportional to $L^{3/2}/h$. Increasing this ratio leads ultimately to a maximum current that the tape cross-section can carry, as ohmic limitation.

Setting tape width $w = 2$ cm leads to mass $m_t = 135$ kg, current $I_a v = 11.9$A, and $power I_{av} E_m L = 2.86$ Mw. For $\beta = 3/4$ and $\beta \approx 0$, we then have $M_{SC} = 560$ kg and 979 kg respectively. Masses, current and power scale linearly with w. In particular, taking $w = 5$ cm leads to masses $m_t = 337.5$ kg, and $M_{SC} = 1400$ kg and 2447 kg for $\beta = 3/4$ and $\beta \approx 3/4$, respectively.

The tether will serve as power source whenever an electric load is plugged in the tether circuit. Note that a power ~ 10 Kw, say, could be substracted from the several Mws produced during capture, with negligible effect on S/C dynamics. Tether current would be off along most of the elliptical orbits, following capture and succesive perijove passes, where the Lorentz force would be weak anyway. The current, however, could be switched on occasionally to generate power.

The orbital energy per unit mass reduced, at capture, in the amount $-(1+\beta)\frac{1}{2}\nu_\infty^2$, determines the semiaxis a_1 of the first elliptical orbit,

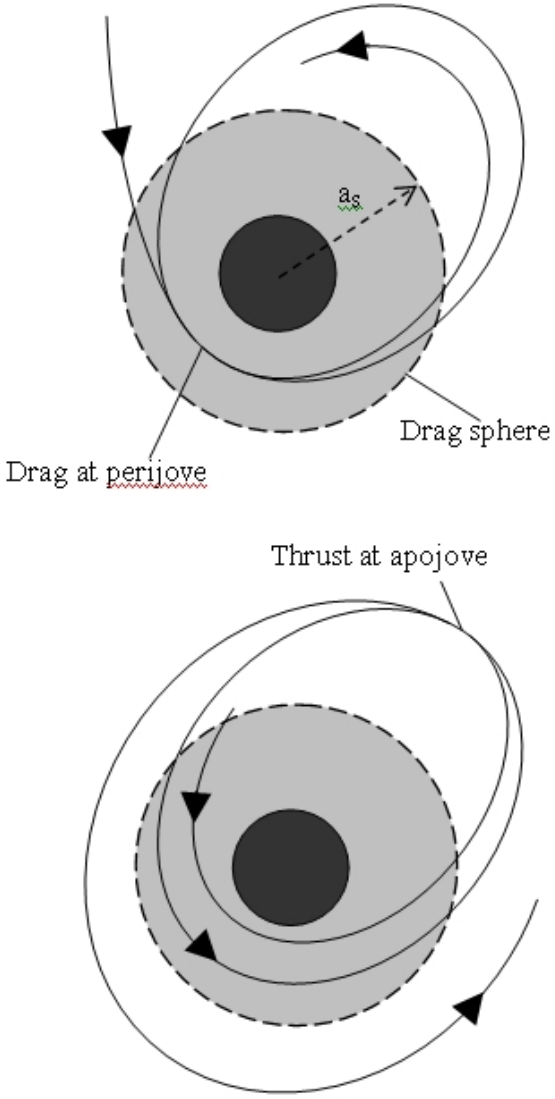


Figure 2: Phases: Capture and lowering apojove. Raising perijove.

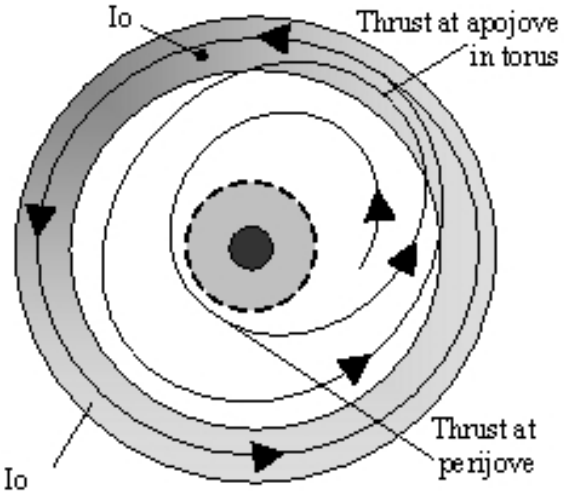


Figure 3: Spacecraft capture at Io

$$-\beta\nu_\infty^2/2 \equiv -\mu_J/2a_1$$

This yields $a_1 = 72.8R_J$, and an orbital period $T_1 = 76.5$ days. Lowering apojove can proceed fast; after the third perijove pass we find $a_3 = 12.8R_J$, $T_3 = 5.6$ days. The raising perijove phase, from $1.5R_J$ to $3R_J$, say, will take somewhat longer.

4. Capture into Low Io Orbit

Assume that in the raising perijove phase, the S/C has been carried to some circular orbit near the end of the plasmasphere, where both B and N_e are sensibly smaller than at capture conditions. Switching tether current on one side of the orbit leads to a sequence of elliptical orbits of fixed perijove but increasing apojove, which can get deep in the *Io* torus after multiple passes. Note that *Io*'s orbital period is only 1.77 days, the periods of those orbits increasing from under to over 1 day, the total duration of the above operation being actually short.

Now the dense, fast-flowing plasma torus can act as a 'filling station'. Switching current on around apojove, tether thrust can take the perijove itself deep into the torus. Finally, with current on and off conveniently, it would be possible to approach *Io* at a small relative velocity to allow it to capture the S/C. Note that the sphere of influence of Io against *Jupiter* is only 7200 km, *Io*'s radius itself being 1820 km. It may be necessary to finely tune tether thrust to keep a low orbit around *Io* stable.

The case for orbiting *Europa* is harder, just meaning that operations should take a sensibly longer time.

5. Conclusions

It appears possible to capture a S/C into orbit around *Jupiter*, and then make the S/C reach a low *Io* orbit, and possibly a low *Europa* orbit, by basically using an ED-tether. Important side issues requiring detailed consideration include tape heating at the intense current collection; S/C survival, under radiation, through the inner Jovian belts and around the moons; and keeping tape dynamics controlled under the low *Jupiter*'s gravity gradient.

Bibliography

[1] F. Bagenal, T. E. Dowling, and W. B. Mc Kinnon, editors. *Jupiter. The Planet, Satellites and Magnetosphere*. Cambridge University Press, Cambridge, UK, 2004. Cambridge Planetary Science.

[2] E. Cahmi and J. Grey. The power potential. *Aerospace America*, 41(8), 2003. editorial.

[3] Space Studies Board. *A science strategy for the exploration of Europa*. National Research Council National Academy Press, Washington, D.C., 1999.

[4] A. Atzei and P. Falkner. Study overview of the jme. Esa technology reference study, European Space Agency, ESTEC-Noordwijk (NL), 2005.

[5] J. R. Sanmartn and E. C. Lorenzini. A 'free-lunch' tour of the jovian system. In *Proc. 8th Spacecraft Charging Technology Conference*, NASA Marshall Space Flight Center, March, 2004.

[6] J. R. Sanmartn and E. C. Lorenzini. Exploration of outer planets using tethers for power and propulsion. *J. Prop. Power*, 21:573–576, 2005.

[7] J. R. Sanmartn, M. Martnez-Sánchez, and E. Ahedo. Bare wire anodes for electrodynamic tethers. *J. Prop. Power*, 9:353–360, 1993.

[8] J. R. Sanmartn. Active charging control and tethers. In *Space Environment Prevention of Risks Related to Charged Particles*, Toulouse, 2002. CNES Space Technology Course. J.P. Catani, Cṕaduès editors.

[9] J. G. Laframboise and L. W. Parker. Probe design for orbit-limited current collection. *Phys. Fluids*, 16:629–636, 1973.

[10] T. W. Hill, A. J. Dessler, and F. C. Michel. Configuration of the jovian magetosphere. *Geophys. Res. Lett.*, 1:3–6, 1974.

[11] N. Divine and H.B. Garrett. Charged particle distributions in jupiter's magnetosphere. *J. Geophys. Res.*, 88:6889–6903, 1983.

Flower Constellations as Rigid Objects in Space

Daniele Mortari

Department of Aerospace Engineering, Texas A&M University, College Station TX, 77843-3141, USA

Abstract

This paper summarizes the findings and the research status on Flower Constellations , a novel and revolutionary way to design satellite constellations that has been discovered and proposed at Texas A&M University. The theory of Flower Constellations is a natural consequence of the theory of compatible (or resonant) orbits. The most surprising aspect of the Flower Constellations is that the satellite distribution identifies the edges of rotating figures whose shapes are time invariant. The complex synchronized dynamics of the satellites preserves the shape of a space object. The whole Flower Constellation is an axial-symmetric rigid object in space that is spinning with prescribed angular velocity. The shape of this object can be deformed by playing with the Flower Constellation design parameters, and the object's axis of symmetry can be set to point to any inertial direction. In particular, when the axis of symmetry is aligned with the Earth's spin axis, the J_2 linear-dominant effect is identical for all the orbits. In this case, the J_2 effect deforms the object shape while preserving the axial-symmetry.

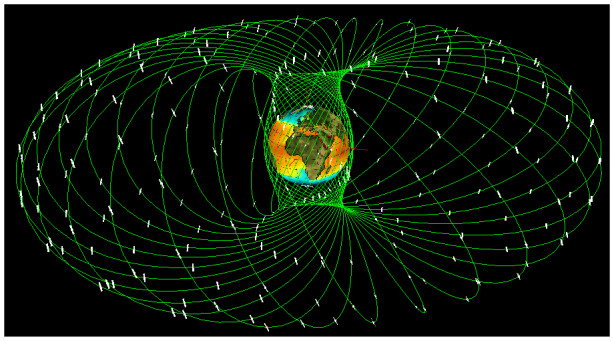
Introduction

The Flower Constellations constitute an infinite set of satellite constellations characterized by periodic dynamics. They have been discovered [1] on the way to the generalization of the concept of some existing satellite constellations. The dynamics of a Flower Constellation identify a set of implicit rotating reference frames on which the satellites follow the same closed-loop relative trajectory [2]. In particular, when one of these rotating reference frames is "Planet Centered, Planet Fixed", then all the orbits become compatible (resonant) with the Planet, and consequently, the projection of the relative trajectory on the planet becomes a repeating ground track.

As a particular case, the Flower Constellations can be designed as J_2 compliant [1, 3], that is with orbit compatibility that takes into account the linear effects of the J_2 perturbation. By considering the J_2 effect on these relative trajectories, it is possible to identify a set of critical inclinations associated with dynamically repeating relative trajectories, called repeating^2 ground track orbits [4], and to identify the two-way orbits [5], having identical and parallel perigee and apogee ground tracks, a property that allows us to design constellations observing the same geographical region from apogee and perigee, simultaneously. The recently proposed *Synodic* and *Relative* Flower Constellations [6, 7, 2] which use dual compatible orbits, as well as the results obtained in designing reconnaissance orbits for Earth sites [8] constitute key initial conditions for many potential research proposals as some of these designs would allow both long-term, stand-off surveillance, and episodic close-in inspection.

In the rotating reference frames the relative trajectories, which depend on five independent integer parameters, constitute a continuous, closed-loop, symmetric pattern of flower petals. Two integer parameters establish the orbit period and the other three distribute the satellites into an upper bounded number of admissible positions. One of the most important consequences of the Flower Constellation theory is that, for a particular set of the five integer parameters, the satellite distribution highlights the existence of *Secondary Paths* [9]. These Secondary Paths, which exhibit many beautiful and intricate dynamics and mysterious properties, are close to being fully understood, and the prediction of them appears to be linked to real algebraic geometry. Finally, the possibilities of re-orienting the Flower Constellation axis and playing with multiple Flower Constellations allow the design of a constellation of constellations, and constellations of formation flying schemes.

The Flower Constellation theory has been developed



at Texas A&M University. Along with the theory, the Flower Constellations Visualization and Analysis Tool (FCVAT) [10] has been developed and coded using Java and Java3D technologies. FCVAT software represents a truly fundamental breakthrough in satellite constellation design methodology, as it makes it easier to see and understand the complicated satellite dynamics, and to see the effects on the constellation of variations of the design parameters. This allows users to easily find different types of satellite formations which have been very difficult to construct using current methods. It is important to emphasize that, in order to design a Flower Constellation , a program like FCVAT must be first developed. Without such a specific program, or equivalent, the design (and the understanding) of Flower Constellation dynamics becomes very difficult or almost impossible.

The Flower Constellations are characterized by an axis of symmetry about which the constellation is rotating in the inertial space as a rigid body and with angular velocity of the rotating compatible reference frame. For Secondary Paths the angular velocity is related to four integer parameters, number of *peri-petals*, *peri-petal step*, number of *apo-petals*, and *apo-petal step* [7, 9].

The dynamics of a Flower Constellation can then be seen as consisting of two distinct parts: (1) an *internal part*, that describes the dynamics of the satellites within the "object-constellation", and (2) an *external part*, where the "rigid object" rotates in inertial space about a spin axis with an angular velocity that can be positive or negative. Some of the resulting shapes are shown in Figs 1 through 2, showing both the versatility and the infinite variety of possible shapes which we call "*choreographes*".

Flower Constellations thus open a new frontier in complex satellite constellations: in particular, these constellation-objects can be used as *building blocks* to construct configurations that can accomplish arbitrarily complex tasks. Indeed, just as the concepts of modularity and functionality gave important paradigm shifts in software design (allowing millions of similar tasks to be treated by one chunk of code), Flower Constellations provide building blocks to enable the creation of arbitrarily complicated ensembles of satellite orbits. Indeed, current approaches to satellites constellation are a simple by-product of the functionality they are designed for. By enabling the research community (and even the general population) to consider constellation as *rigid objects*, we enable new functionalities of satellites in urgently needed applications, and the study of even more intricate constellations for which functionalities have yet to be found.

Flower Constellations also allow us to profitably transform our intuition by thinking of trajectories in the solar system not just as shells like the LEO/MEO/GEO elliptical orbits, but rather as the union of several objects represented by different Flower Constellations . In interferometry, for example, a star-like Flower Constellation (see Fig. 17) can be thought as a unique radar-like antenna instead of as a collection of spacecraft.

Other interesting Flower Constellations are characterized by morphing capabilities, as for the morphing Flower Constellation shown in Figs. 3 and 4. This constellation has a dynamics that periodically changes from a five-loop aspect (Fig. 3) to an inscribed pentagon aspect (Fig. 4).

The particular dynamics of a Flower Constellation are obtained by introducing an automatic mechanism, ruled by a set of three integer parameters, to distribute the satellites into a limited set of "admissible locations". This is shown in Fig. 7, where 17 spacecraft are located on the same inertial orbit (green) and all of them belong to the same ECEF relative trajectory (red). These parameters rule the important phasing of the Flower Constellations . In this way, this new methodology to design satellite constellations has greatly simplified the constellation design problem and, thus, has provided the means to solve an extremely difficult family of problems.

Recently, two novel constellation design methodologies have been proposed [6, 2]. These are the Synodic and the Dual (or Relative) Flower Constellations , which constitute the important extension of Flower Constellations synchronized with the motion of two celestial objects (e.g., two planets) orbiting about the same gravitational mass. These two rotating reference frames can also be associated with natural or artificial satellites (e.g., moons, spacecraft) orbiting about a planet, and one of these frames can also be associated with the rotation of the central body itself. In particular, a Synodic Flower Constellation is made with orbits that are compatible with a reference frame rotating with a period suitably derived from the synodic period of the two objects, while a Dual Flower Constellation is made of orbits that are,

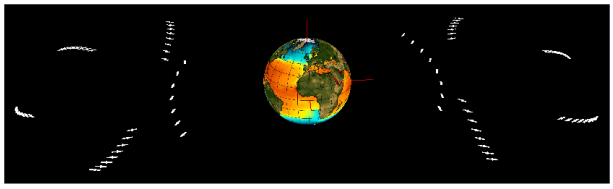


Figure 1: King Defense Flower Constellation

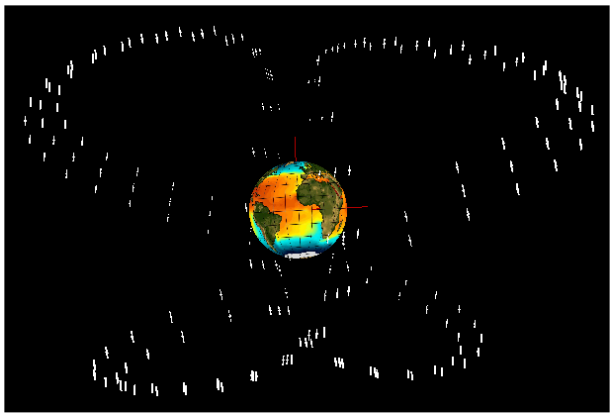


Figure 2: Two 3×3 formation flying loops

simultaneously, compatible with both the objects rotating reference frames. The latter, however, can be achieved under a very particular condition, that can be numerically approximated. The resulting constellation dynamics is synchronized with the dynamics of the geometrical rotation of the two objects.
The Flower Constellations , with its latest "*Synodic*" and "*Relative*" extensions, have already been partially investigated and some results for classical applications been obtained.

Compatible orbits

The Flower Constellation theory is built and derived from the theory of compatible (or resonant) orbits. The "compatibility" is a synchronization property between two rotating reference frames. Mathematically, the rotating reference frames, $\mathcal{F}_1$ and $\mathcal{F}_2$, are *compatible* or *resonant*, if their constant angular velocities, ω_1 and ω_2, satisfy the relationship

$$N_1\,\omega_2 = N_2\,\omega_1 \tag{1}$$

where N_1 and N_2 can be any integers. In the case the angular velocities are not constant, then $\mathcal{F}_1$ and $\mathcal{F}_2$ are compatible iff $\omega_1(t)$ and $\omega_2(t)$ are periodic functions. In this case, the compatibility is specified by the relationship

$$N_1 \int_0^{T_2} \omega_2\,dt = N_2 \int_0^{T_1} \omega_1\,dt \tag{2}$$

where T_1 and T_2 are the periods of the rotating frames.
The orbit mean motion n is a fictitious constant angular velocity associated with the periodic motion of the satellite along its orbit. Therefore, an orbit is compatible, with respect to a reference frame $\mathcal{F}$ rotating with angular velocity ω, if the orbit period T satisfies the relationship

$$N_p\,T = N_p\,\frac{2\pi}{n} = N_d\,\frac{2\pi}{\omega} \tag{3}$$

where N_p and N_d are two positive integers characterizing compatible orbits. Alternatively, the definition of a compatible orbit can be expressed by saying that an orbit is compatible when the ratio of its period with that of the rotating reference frame is rational.
Equation (3) simply states that after N_p orbital periods the rotating reference frame has performed N_d complete rotations and, consequently, the satellite and the rotating reference frame come back to their initial positions. This implies that the trajectory of the satellite in the rotating reference frame - the *relative trajectory* - is a continuous closed-loop trajectory that can be seen as a closed-loop 3-dimensional space track.[1] Two examples are provided in Figs. 5 and 6. In particular, $N_p\,T$ represents the time required by the satellite to repeat the entire relative trajectory.
In the Flower Constellation theory the two integers, N_p and N_d, are identified as the *Number of Petals* and the *Number of Days*, respectively. The reason for that arises because if the rotating reference frame is selected to be

[1] Even though compatible orbits are known as "repeating ground track" orbits, we want to highlight the distinction between these two definitions. A ground track is just the projection of the closed-loop trajectory on the Earth surface. This projection does not contain the full information of the 3-D trajectory. Moreover, the set of compatible orbits is just a subset of the repeating ground track orbits set (e.g., any equatorial orbit is repeating ground track!).

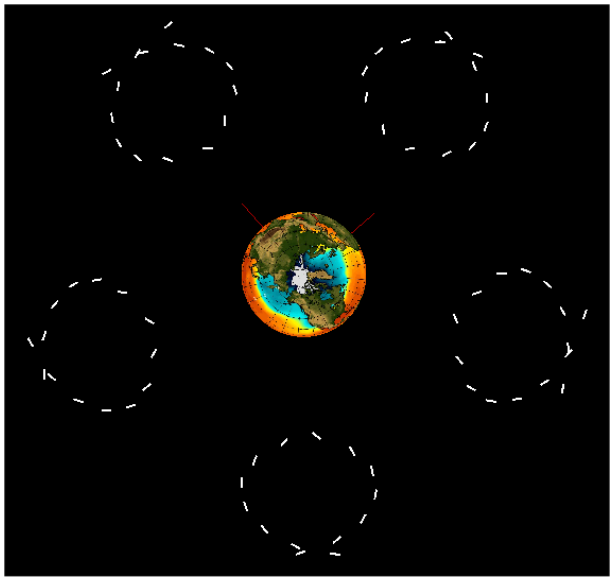


Figure 3: Morphing Flower Constellation : Five-loop

Earth-Centered Earth-Fixed, then N_d really represents the number of days to repeat the relative trajectory, while the *Number of Petals* N_p, that actually represent the number of orbit revolutions, finds its origin because of the petal-like shape of the relative trajectory.
It is important to understand that an orbit, that is compatible with respect to an assigned rotating reference frame, is also compatible with an infinite set of rotating reference frames. In fact, an orbit satisfying Eq. (3) is also compatible with all the reference frames $\mathcal{F}'$ rotating with angular velocity

$$\omega' = \omega \left(\frac{N'_d}{N'_p}\right) \left(\frac{N_p}{N_d}\right) \tag{4}$$

where N'_p, N'_d, and ω', satisfy the compatibility condition

$$N'_p \, T = N'_d \, \frac{2\pi}{\omega'} \tag{5}$$

Flower Constellation Phasing

Flower Constellations are built with the constraint that all the satellites belong to the same relative trajectory. In order to obtain the mathematical relationship stating this property, let us consider two identical compatible orbits having node lines displaced from each other by $\Delta\Omega$. Let us consider, as initial condition ($t = 0$), the satellite in the first orbit be located at pericenter ($M = 0$). This implies that the time interval Δt spent by the rotating reference frame to rotate of $\Delta\Omega$ (the relative trajectory is fixed in the rotating reference frame) must be identical to that associated with the increase of the mean anomaly of the spacecraft along its orbit. Therefore, we can write the relationship

$$\Delta t = -\frac{\Delta\Omega}{\omega} = \frac{\Delta M}{n} \tag{6}$$

where ω is the angular velocity of the rotating reference frame and n the orbit mean motion. The reason of the negative sign depends on the fact that for positive ω, the inertial orbits rotate clock-wise while Ω increases counter clock-wise. Figure 1 of Ref. [1] helps in understanding it. Equation (6) shows a direct relationship between right ascension of the ascending node and mean anomaly. Substituting Eq. (3) in Eq. (6) we obtain

$$-N_p \, \Delta\Omega = N_d \, \Delta M \tag{7}$$

This relationship, which represents the fundamental equation of the Flower Constellation phasing, allows us to evaluate the "*admissible locations*" where to place the constellation satellites in order they all belong to the same relative trajectory. In other words, if a satellite is located at position M_1 of the orbit characterized by Ω_1 then, in order to belong to the same relative trajectory, a satellite on a different orbit characterized by Ω_2 must be placed at position M_2, where

$$M_2 = M_1 - (\Omega_2 - \Omega_1) \, \frac{N_p}{N_d} \tag{8}$$

In the case the second orbit coincides with the first orbit ($\Omega_2 - \Omega_1 = F_h \, 2\pi$, where F_h can be any integer), then Eq. (8) highlights all the admissible locations per orbit

$$M_{1,k+1} = M_{1,k} - F_h \, 2\pi \, \frac{N_p}{N_d} \tag{9}$$

where $F_h = 0, 1, \cdots, N_d - 1$. Equation (9) allows us to state the following:

1. Two satellites on the same orbit and displaced by $\Delta M = 2\pi \, F_h \, N_p / N_d$, where F_h can be any integer, belong to the same relative trajectory.

2. The number of admissible locations per orbit in Flower Constellation is N_d.

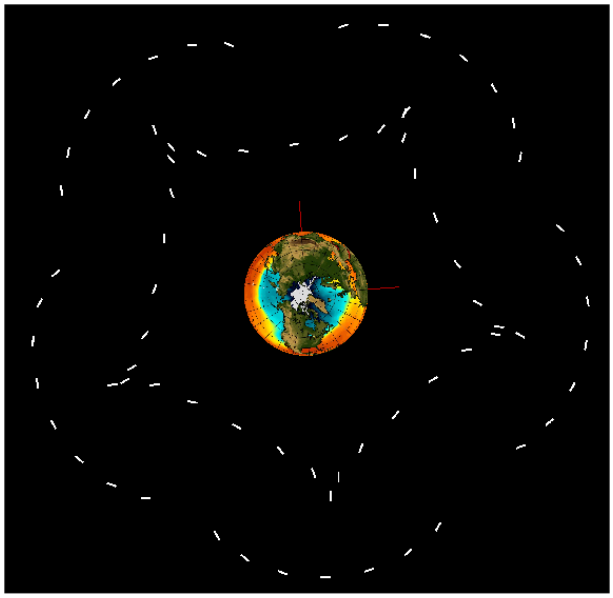


Figure 4: Morphing Flower Constellation : Inscribed pentagon

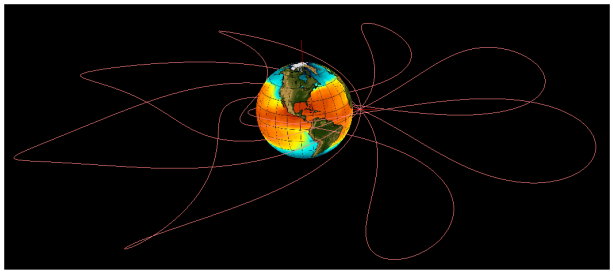


Figure 5: Relative trajectory example #1

Equation (8) provides us with the *natural* admissible location where to place the satellite. However, since there are N_d admissible locations per orbit, all the admissible locations in the orbit characterized by Ω_2 are provided by the relationship

$$M_2 = M_1 - (\Omega_2 - \Omega_1)\,\frac{N_p}{N_d} - F_h\,2\pi\,\frac{N_p}{N_d} \tag{10}$$

Consequently: we have the complete free choice of where to place the first satellite (Ω_1, M_1), but when this is done then, for any assigned number N of orbits (not necessarily evenly distributed in 2π), the admissible locations are all defined. It is clear that evenly orbits distribution are preferred whenever the symmetry is desired. In the Flower Constellations this is obtained by selecting a rational value of 2π for the orbit node lines step

$$\Delta\Omega = \Omega_{k+1} - \Omega_k = 2\pi\,\frac{F_n}{F_d} \tag{11}$$

where F_n and F_d can be any two integers. If the first orbit is selected having $\Omega_1 = M_1 = 0$, then Eq. (11) provides us with the orbit node lines sequence

$$\Omega_k = 2\pi\,\frac{F_n}{F_d}\,(k-1) \tag{12}$$

while Eq. (8) the associated admissible locations

$$M_k = 2\pi\,\frac{F_n\,N_p + F_d\,F_h}{F_d\,N_d}\,(1-k) \tag{13}$$

Equation (13) governs the sequence of the mean anomalies, which is dictated by the rational parameter $(F_n\,N_p + F_d\,F_h)/(F_d\,N_d)$. This ratio might be further simplified. To this end, let $C = \gcd(F_n\,N_p + F_d\,F_h, F_d\,N_d)$. This implies that Eq. (13) can be re-written in the following simplified way

$$M_k = 2\pi\frac{F_nN_p + F_dF_h}{F_dN_d}(1-k) = 2\pi\frac{R_n}{R_d}(1-k) \tag{14}$$

where $k = 1, 2, \ldots, N_s$, and where

$$R_n = \frac{F_n\,N_p + F_d\,F_h}{C} \quad \text{and} \quad R_d = \frac{F_d\,N_d}{C} \tag{15}$$

Equation (14) implies that when we come back to the initial orbit with the sequence index $k = F_d + 1$, then the mean anomaly of the second satellite belonging to the first inertial orbit is

$$M_{F_d+1} = -2\pi\,\frac{R_n}{R_d}\,F_d \tag{16}$$

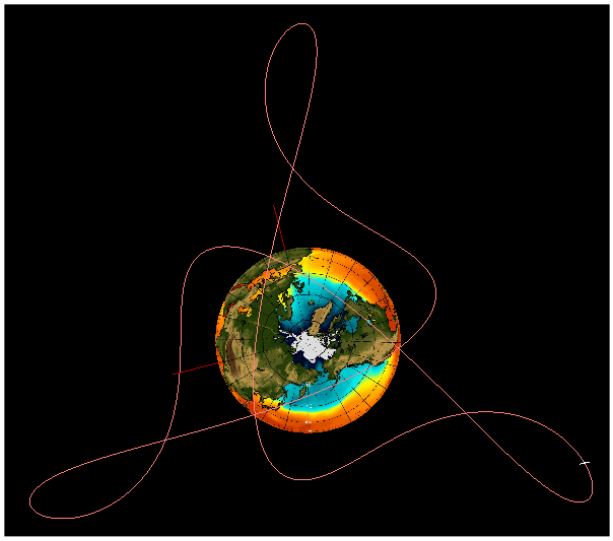


Figure 6: Relative trajectory example #2

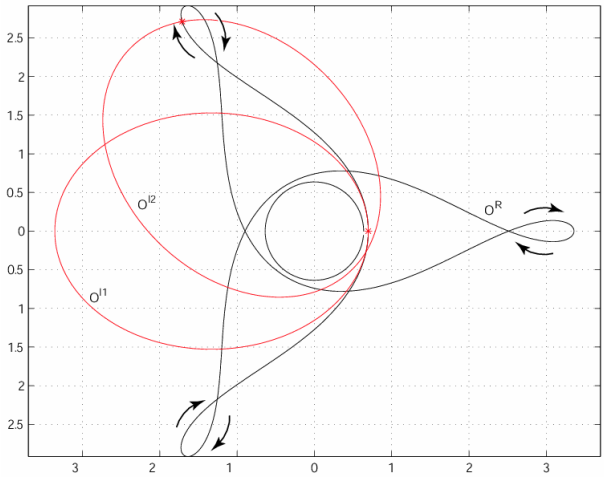


Figure 7: Phasing Geomery

Let $C_r = \gcd(F_d, R_d)$. Therefore, the integer

$$N_{so} = \frac{R_d}{C_r} \qquad \text{where} \qquad N_{so} \leq N_d \tag{17}$$

represents the number of satellites per orbit for the chosen distribution sequence and constellation. As a consequence, the total number of satellites will be

$$N_s = N_{so}\, F_d \leq N_d\, F_d \tag{18}$$

The parameter N_{so} also represents the number of loops around the Earth that are completed while placing satellites in a Flower Constellation . That is to say, if one places a single satellite in each of F_d orbits, it will take you N_{so} cycles to place all the satellites. Now, since N_d represents the overall number of admissible locations in one orbit, then N_{so} tells you how many of the N_d locations are filled in a given satellite distribution. Therefore, if $N_{so} = N_d$ then all the available admissible spots are filled, while if $N_{so} = 1$ then only one admissible spot (per orbit) is used. If $N_{so} < N_d$, then we define $N_s \equiv N_{so}\, F_d$ and describe the satellite distribution as forming a N_{so}/N_d *Secondary Path*.

In summary, associated with a given distribution sequence there is always an upper limit for the number of admissible locations where one can locate satellites. Therefore, a single Flower Constellation cannot be host to more than N_s satellites, where

$$N_s \leq N_d\, F_d \tag{19}$$

However, for an assigned sequence distribution, there exists the possibility that the sequence distribution does not fill all the $N_d\, F_d$ admissible locations. This happens when, during the satellite distribution, a satellite should be placed onto the initial location of the first satellite, that has already been occupied. When this happens the satellites are distributed along a *secondary path*, which is associated with a sequence distribution that creates a

premature closing loop. Depending upon the number of satellites per orbit N_{so} constituting this particular distribution, a classification of the secondary paths is given. So, a secondary path having N_{so} satellites/orbit is called "secondary path of order N_{so}".

Secondary Paths

When all the admissible locations of a Flower Constellation are filled (especially when the number of these locations are many), the Flower Constellation dynamics reveals the shape of the relative trajectory by clearly showing the number of petals (apogees of the relative trajectory). In this case the whole constellation appears to be rotating, as a rigid body, with the angular velocity of the planet, if the orbits are compatible with the planet's spin rate. Sometimes, however, the phasing does not allow us to fill out all the admissible locations and it happens that the satellite distribution sequence comes back to the first position ($\Omega = 0$ and $M = 0$) before all the admissible locations are filled. When this happens, the Flower Constellation dynamics highlights the existence of *Secondary Paths* (SP) that have unexpected and beautiful shapes that are time invariant [3].
The immobility of the printed figure does not allow us to demonstrate the resulting complex shape-preserving dynamic. While complete Flower Constellations spin with a prescribed angular velocity (i.e. the same rate as that of the rotating reference frame), the spin rate of a secondary path should be quantified. Note that the angular velocity of a secondary path is *apparent* and not real. That is to say, the apparent angular rotation is not a motion that can be described by any particular dynamical relationship but rather is an artifact of the mathematics that generates a Flower Constellation . In other words, the appearing angular rotation IS NOT continuous but *appears* continuous. However, the continuity nor is discrete, as in the effect of the fast flow of photograms of motion pictures, because the satellites motion IS continuous. In effect, the angular motion pops up because of a particular combination of the continuity motion of a satellite along its orbit and the discrete separation of contiguous orbits.

Loops, petals, and jumping parameters

While in a complete Flower Constellation the satellites highlight the shape of the relative trajectory by moving along the single loop, in secondary paths the satellites can form single ($N_\ell = 1$) or multiple ($N_\ell > 1$) loops[2]. Figures (8) and (9) show a single- and a double-loop Secondary Paths, respectively. In the following, we introduce and explain the parameters characterizing the Secondary Paths and the relationships between them.
In addition to the number of loops N_ℓ, a Secondary Path is characterized by four integer parameters: the overall number of apogees (apo-petals), $N_{a\ell}$, the overall number of perigees (peri-petals), $N_{p\ell}$, and the two jumping-petal step parameters, $J_{a\ell}$ and $J_{p\ell}$, indicating the petal sequence visited by any satellite while moving from the petal k to the petal $(k + J_{a\ell})$ or $(k + J_{p\ell})$, where the apo/peri-petals are counted counter clockwise, and where $0 \leq J_{a\ell} < N_{a\ell}$ and $0 \leq J_{p\ell} < N_{p\ell}$.
Each loop is characterized by $N_{a\ell}/N_\ell$ apo-petals and by $N_{p\ell}/N_\ell$ peri-petals. Therefore, in a single loop the angles between any two consecutive apo-petals and between any two consecutive peri-petals are $2\pi\, N_\ell/N_{a\ell}$ and $2\pi\, N_\ell/N_{p\ell}$, respectively.
Most of the Secondary Paths are characterized by $N_{a\ell} = N_{p\ell}$. However, for some sets of design parameters, it is possible to have the number of apo-petals different from the number of peri-petals. As example, Fig. (10) shows a Secondary Path having $N_{a\ell} = 10$ and $N_{p\ell} = 5$.
In a Secondary Path the time required for a satellite to move from a petal to the next is, clearly, one orbit period. Therefore, in order to complete the loop, the value of the jumping parameter must be consistent with the number of petals. This consistency is mathematically defined by the following property: the greatest common divisor between the jumping parameter and the number of petals must be one for the apo-petals

$$\gcd(N_{a\ell}/N_\ell, J_{a\ell}) = 1 \tag{20}$$

and one for the peri-petals

$$\gcd(N_{p\ell}/N_\ell, J_{p\ell}) = 1 \tag{21}$$

This propriety assures the connection between all the petals of a loop, that is, it ensures that the satellites visit all the petals of the loop to which they belong.
After one orbit period, the satellite comes back to its initial position, but on a different petal of its loop[3]. After $N_{a\ell}$ orbit periods, the satellite has completed the visiting of all the apo-petals of his loop. In the case the Secondary Path has just an $N_\ell = 1$ single loop, $N_{a\ell} = 2$ apo-petals, and an $J_{a\ell} = 1$ jumping apo-petal step parameter, then after $N_{a\ell}$ orbit periods the Secondary Path is rotated by an angle 2π and, therefore, the Secondary Path angular velocity is $\omega = 2\pi/(N_{a\ell}\, T)$.
The expression for the general case, when the Secondary Path loop is characterized by any value of $N_{a\ell}$ and $J_{a\ell}$, can be easily derived. For the apo-petals we have two distinct solutions. One is associated with a clockwise loop rotation

$$\omega_{a\circlearrowright} = \frac{2\pi}{T} \cdot \frac{N_{a\ell} - J_{a\ell}}{N_{a\ell}} \tag{22}$$

while the second is associated with counter clockwise loop rotation

$$\omega_{a\circlearrowleft} = -\frac{2\pi}{T} \cdot \frac{J_{a\ell}}{N_{a\ell}} \tag{23}$$

[2]The shape of the loops in multi-loop Secondary Paths are all identical. They are just rotated one to another by an angle $2\pi/N_\ell$.

[3]If it comes back on the same petal, then the problem becomes trivial because it implies that the angular velocity of the Secondary Path is identical to that of the rotating reference frame. This cannot happen in Secondary Path but just in complete Flower Constellations .

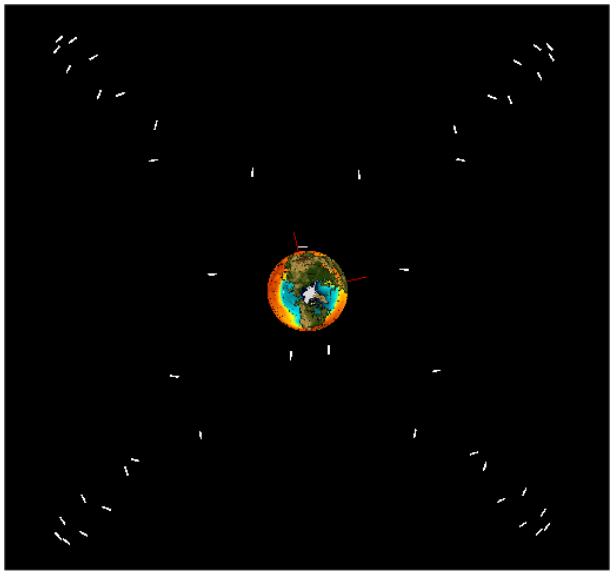


Figure 8: Single-loop Secondary Path

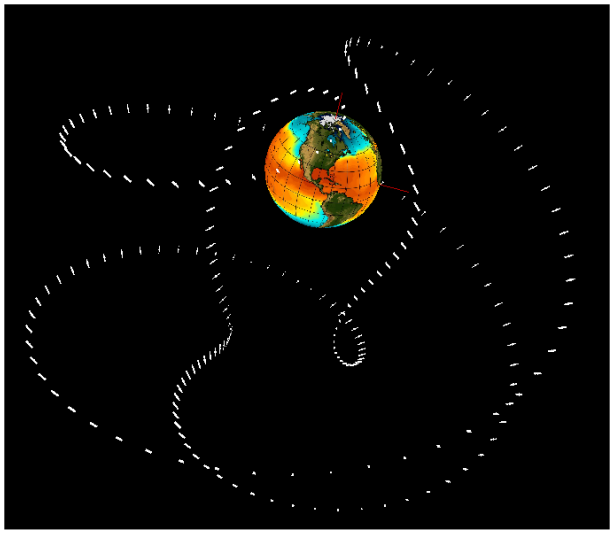


Figure 9: Double-loop Secondary Path

Analogously, for the peri-petals and in the general case when $N_{a\ell} \neq N_{p\ell}$ (and $J_{a\ell} \neq J_{p\ell}$), we have two distinct angular velocities: the clockwise

$$\omega_{p\circlearrowleft} = \frac{2\pi}{T} \cdot \frac{N_{p\ell} - J_{p\ell}}{N_{p\ell}} \tag{24}$$

and the counter clockwise

$$\omega_{p\circlearrowright} = -\frac{2\pi}{T} \cdot \frac{J_{p\ell}}{N_{p\ell}} \tag{25}$$

respectively. Equations (22) through (25) show that the angular velocity of a Secondary Path does not depend on N_ℓ,

Flower Constellation orientation

A non-oriented Flower Constellation has the characteristic property of having the axis of symmetry coincident with the planet's spin axis. The main reason is because two important orbital parameters - inclination and right ascension of the ascending node - are derived with respect to that axis. In the important case of choosing the constellation axis of symmetry as the planet's spin axis and the rotating reference as "planet-centered planet-fixed", then all the Flower Constellation satellites will travel along an identical repeating ground track. However, in general, the pointing of the axis of symmetry of a Flower Constellation is (as for the angular velocity of the reference rotating frame) a choice that is left completely free to users.

When choosing the constellation's axis of symmetry to

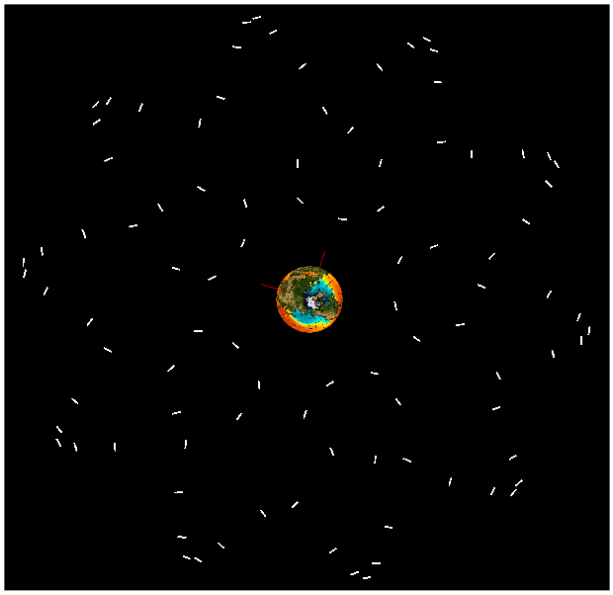


Figure 10: SP with $N_{a\ell} = 10$ and $N_{p\ell} = 5$

not be coincident with the planet's spin axis, then it is important to be aware that all the orbits of the constellation will have, in general, different inclination and right ascension of the ascending node. This implies that each orbit is subjected to different J_2 perturbations. Therefore, the deformation of the relative trajectory will be different for each orbit and, consequently, the beautiful symmetrical dynamics will be destroyed, unless using active control to compensate the relative perturbations and maintain the constellation dynamics.

To evaluate inclination and right ascension of ascending node of an oriented Flower Constellation we proceed as follows. Let **r** and **v** be the position and velocity inertial vectors (cartesian coordinates), respectively, of a generic satellite on a non-oriented Flower Constellation . In particular, **r** and **v** can be expressed in term of orbital parameters

$$\mathbf{r} = R_{OI}^{\mathrm{T}} \frac{p}{1 + e\cos\varphi} \left\{ \begin{array}{c} \cos\varphi \\ \sin\varphi \\ 0 \end{array} \right\} \tag{26}$$

and

$$\mathbf{v} = R_{OI}^{\mathrm{T}} \sqrt{\frac{\mu}{p}} \left\{ \begin{array}{c} -\sin\varphi \\ e + \cos\varphi \\ 0 \end{array} \right\} \tag{27}$$

where e is the orbit eccentricity, p the semilatus rectum, φ the true anomaly, and

$$R_{OI} = R_3(\omega)\, R_1(i)\, R_3(\Omega) \tag{28}$$

is the orthogonal transformation matrix moving from Inertial to Orbital reference frame. Matrices $R_1(\vartheta)$ and $R_3(\vartheta)$ are the matrices performing rigid rotation about the first and third coordinate axis

$$R_1(\vartheta) = \left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & \cos\vartheta & \sin\vartheta \\ 0 & -\sin\vartheta & \cos\vartheta \end{array} \right] \tag{29}$$

and

$$R_3(\vartheta) = \left[\begin{array}{ccc} \cos\vartheta & \sin\vartheta & 0 \\ -\sin\vartheta & \cos\vartheta & 0 \\ 0 & 0 & 1 \end{array} \right] \tag{30}$$

Now, let

$$\hat{\mathbf{d}} = \left\{ \begin{array}{c} \sin\alpha\,\cos\beta \\ \sin\alpha\,\sin\beta \\ \cos\alpha \end{array} \right\} \tag{31}$$

be the direction of the desired Flower Constellation axis (where α and β are colatitude and longitude of the Flower Constellation axis in ECI). This implies that all the orbits of the Flower Constellation must be rotated by the angle α about the axis

$$\hat{\mathbf{a}} = \left\{ \begin{array}{c} -\sin\beta \\ \cos\beta \\ 0 \end{array} \right\} \tag{32}$$

The matrix performing such a rigid rotation is

$$R(\hat{\mathbf{a}}, \alpha) = I_3\, \cos\alpha + (1 - \cos\alpha)\, \hat{\mathbf{a}}\hat{\mathbf{a}}^{\mathrm{T}} + \tilde{A}\, \sin\alpha \tag{33}$$

where I_3 is the 3×3 identity matrix and

$$\tilde{A} = \left[\begin{array}{ccc} 0 & 0 & \cos\beta \\ 0 & 0 & \sin\beta \\ -\cos\beta & -\sin\beta & 0 \end{array} \right] \tag{34}$$

is the skew-symmetric matrix performing the vector cross-product.

Now, the rotated orbit has, in general, new values for inclination, argument of perigee, and right ascension of

ascending node that can be derived from the new rotated cartesian vectors

$$\mathbf{r}_n = R(\hat{\mathbf{a}}, \alpha)\,\mathbf{r} \qquad \text{and} \qquad \mathbf{v}_n = R(\hat{\mathbf{a}}, \alpha)\,\mathbf{v} \tag{35}$$

using well known transformations.

Dual-Compatible Flower Constellations

This section analyzes some particular Flower Constellations whose orbits are simultaneously compatible with two rotating reference frames. As it will be demonstrated later, for these Flower Constellations , we are no longer free to choose where to locate the orbit apsidal lines (i.e the values of Ω_k). Furthermore, the overall number of admissible locations strongly depend on the design parameters.

An orbit is Dual-Compatible (or dual-resonant) if, assigned the four integers N_{p1}, N_{d1}, N_{p2}, and N_{d2}, its orbital period T satisfies the two relationships

$$N_{p1}\,T = N_{d1}\,T_1 \qquad \text{and} \qquad N_{p2}\,T = N_{d2}\,T_2 \tag{36}$$

where

$$T_1 = \frac{2\pi}{\omega_1} \qquad \text{and} \qquad T_2 = \frac{2\pi}{\omega_2} \tag{37}$$

are the periods associated with two reference frames rotating with angular velocities ω_1 and ω_2, respectively.

Based on the above definition, any orbit characterized by orbit period T, is compatible with an infinity of rotating reference frames characterized by the set of angular velocities

$$\omega_k = \frac{2\pi}{T}\,\frac{N_{dk}}{N_{pk}} = n\,\frac{N_{dk}}{N_{pk}} \tag{38}$$

where n is the orbit mean motion.

In order to find out where to locate the satellite of a Dual-Flower Constellation , let us evaluate the RAAN variation between two consecutive satellites

$$\Delta\Omega = \Omega_{k+1} - \Omega_k \tag{39}$$

where to allocate one orbit $(k+1)$ with respect to the previous one (k). The satellite in the $(k+1)$-th orbit will have a variation of the mean anomaly with respect to the value of the previous satellite that can be evaluated using Eq. (10). The ΔM expression is

$$\Delta M = M_{k+1} - M_k = -\Delta\Omega\,\frac{N_p}{N_d} - 2\pi\,\frac{F_h}{N_d} \tag{40}$$

We can evaluate the variation ΔM using both sequences, $\mathfrak{F}_1 \triangleq \{N_{p1}, N_{d1}, F_n, F_d, F_h\}$ and $\mathfrak{F}_2 \triangleq \{N_{p2}, N_{d2}, F_n, F_d, F_h\}$. These two distributions, $\mathfrak{F}_1$ and $\mathfrak{F}_2$, must provide values for ΔM that can differ just of $2\pi\,\ell$, where ℓ can be any integer. Therefore, we can write that

$$\begin{aligned} \Delta M = &\quad -\Delta\Omega\,\frac{N_{p1}}{N_{d1}} - 2\pi\,\frac{F_h}{N_{d1}} = \\ = &\quad -\Delta\Omega\,\frac{N_{p2}}{N_{d2}} - 2\pi\,\frac{F_h}{N_{d2}} + 2\pi\,\ell \end{aligned} \tag{41}$$

This equation allows us to obtain an expression for $\Delta\Omega$ that is a function of the integer parameter ℓ

$$\begin{aligned} \Delta\Omega_\ell = &\quad 2\pi\,\frac{F_h(N_{d2} - N_{d1}) + \ell\,N_{d1}\,N_{d2}}{N_{d1}\,N_{p2} - N_{d2}\,N_{p1}} = \\ = &\quad 2\pi\,\frac{G_\ell}{G_d} \end{aligned} \tag{42}$$

which allows us to evaluate the values (therefore, the sequence) of the right ascension of the ascending nodes, Ω_k, where the two distributions locate satellites with the same values of the mean anomaly (same orbital position). In order to use Eq. (42), the condition $G_d \neq 0$, that is

$$N_{d1}\,N_{p2} \neq N_{d2}\,N_{p1} \tag{43}$$

must be satisfied. This condition implies that the case $T_1 = T_2$ should be avoided. In addition, the values of ℓ satisfying

$$\ell\,N_{d1}\,N_{d2} = F_h(N_{d1} - N_{d2}) \quad \leftrightarrow \quad G_\ell = 0 \tag{44}$$

which are associated with the condition $\Delta\Omega_\ell = 0$, allow us to obtain the sequence of all the solutions per orbit. This sequence can also be obtained for the values of ℓ giving $\Delta\Omega_\ell = 2\pi\,m$, which is satisfied when

$$\gcd\,(G_\ell, G_d) = G_d \tag{45}$$

For each value of $\Delta\Omega_\ell$ provided by Eq. (42), we have an associated value for ΔM_ℓ

$$\begin{aligned} \Delta M_\ell &= -\Delta\Omega_\ell\,\frac{N_{p1}}{N_{d1}} - 2\pi\,\frac{F_h}{N_{d1}} = \\ &= -\Delta\Omega_\ell\,\frac{N_{p2}}{N_{d2}} - 2\pi\,\frac{F_h}{N_{d2}} + 2\pi\,\ell \end{aligned} \tag{46}$$

Summarizing, a dual-compatible Flower Constellation is built using the phasing sequence

$$\begin{cases} \Omega_{k+1} = \Omega_k + \Delta\Omega_\ell \\ M_{k+1} = M_k + \Delta M_\ell \end{cases} \tag{47}$$

where $k = 1, 2, \cdots$, and $\Delta\Omega_\ell$ and ΔM_ℓ are provided by Eqs. (42) and (46), respectively.

Examples and Potential Applications

The Flower Constellations and the recently introduced Dual Flower Constellations combine a number of new attractive features suitable for many potential classic applications (communications, Earth and deep space observation, coverage, navigation systems, etc.), as well as for new and advanced concepts. Some Flower Constellations schemes can be suitable for very futuristic applications while other can be of immediate use. Let us briefly describe soem of these potential applications:

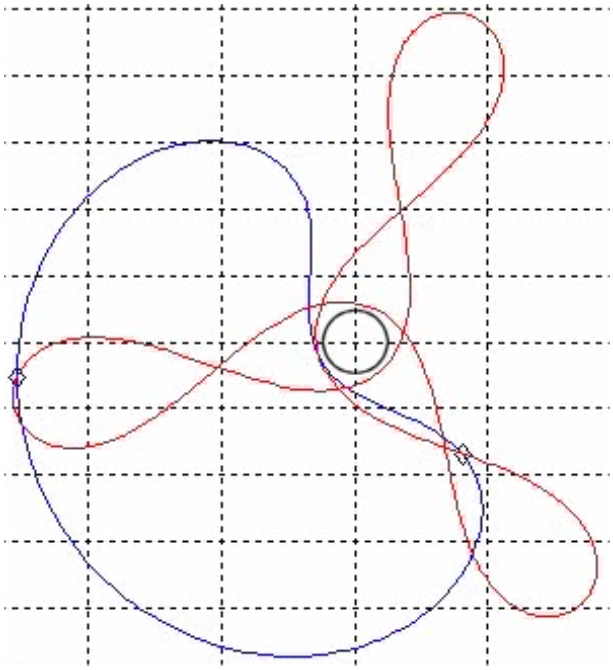


Figure 11: Dual Flower Constellation Example #1

1. *Space Network Architecture* (SNA) for planetary communications. Interplanetary communications are presently performed by means of single-hop links. In this simple architecture there is one node at the exploration planet (e.g. Mars) and one node at the Earth (specifically, the antennae of the NASA Deep Space Network). This simple architecture presents two severe constraints: it requires direct visibility (and hence limited duration operation) and it does not tolerate node failure. Using Dual Flower Constellations we can design a constellations that are synchronized with the motion of two rotating reference frames (e.g., Earth and planet orbital periods). Dual Flower Constellations [6, 7, 2] can provide solutions that avoid the mentioned critical constraints and would improve the communications necessary for human planetary missions. The design of a SNA using Dual Flower Constellation could potentially consist of multi-hop links, a constellation of spacecrafts connecting the Earth with a mission planet (or moon) and would drive to improve the connectivity of the deep space network. Reference [6] introduced this idea and proposed some approximated solutions to help communication for future missions to Mars and/or Jupiter.

   Reference [7] provided novel insights on the theory (specially on the phasing rules) while the complete mathematical theory (tractatus) on Flower Constellations will be presented in Refs. [2] and [9]. Figure 11 and 12 show two Dual Flower Constellation examples. In these figures the relative trajectories, associated with two distinct and independent reference frames are provided. The spacecrafts are located at some "admissible" intersections of the two relative trajectories that rotates with different constant angular velocities.

   The design of a *Space Network Architecture* for planetary communications must take into proper consideration the effects of the orbital geometry on the network topology, and the resulting effects of path delay and handover on network traffic (due to the great distances involved). In addition to these problems, a wide variety of requirements and constraints must also be satisfied. These are:

   (a) *Service continuity*: if any one of the nodes becomes inoperative (either, permanently or momentarily), then the communications are still guaranteed.

   (b) *Power efficiency*: minimize inter-node angle variations (to narrow the antennae FOV), minimize inter-node distances (to limit communication power), etc,

   (c) *Time efficiency*: minimize the overall distance (to limit communication times), and

   (d) *Fuel efficiency*: seek to minimize the orbit maintenance requirements by optimizing amongst feasible orbit configurations.

2. *Solar Global Navigation System*. This can be investigated using Flower Constellations synchronized with a reference frame rotating with the Earth orbit mean motion. The existing

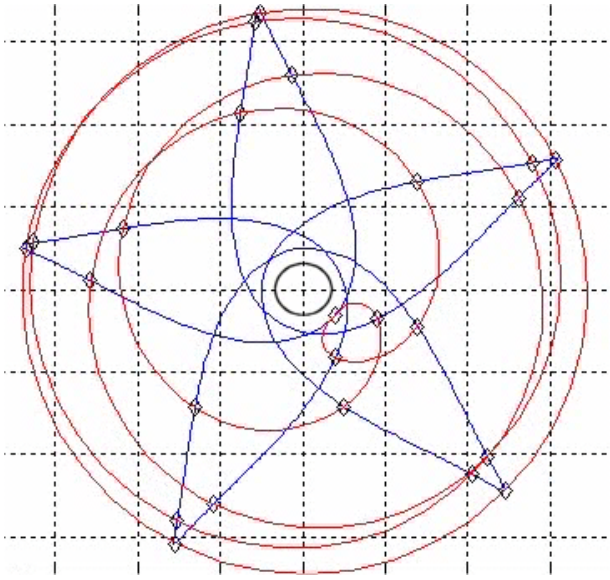


Figure 12: Dual Flower Constellation Example #2

Global Navigation systems (GPS, GLONASS, GalileoSat) are build using circular orbits (Walker constellations), only. European GalileoSat constellation is designed as in Fig. 13 with satellites lying on three orthogonal orbit planes.

This choice creates eight holes, one for each octant, which keep the satellites allocation far from being uniformly distributed in space. A first attempt to design a Global Navigation Flower Constellation (GNFC), has brought to the solution scheme shown in Fig. 14 [11, 12], where the optimality is defined as the most uniform satellite distribution along the relative trajectory. Using the same number of satellites as GalileoSat, GNFC provides better Geometric *and* Attitude Dilution of Precision parameters [11, 12] or the same level of accuracy with lesser satellites.[4]

3. *Space Dynamo*. The Faraday law of inductance states that a voltage is generated by a coil of wire when the magnetic flux enclosed by it changes. A space dynamo, for energy production in space, can be obtained using Flower Constellations with multiple Secondary Paths forming inclined circles. In this configuration, each Secondary Path can be considered a very long single wire where the circuit could be closed by electron cannons. By orbiting, each wire experiences the variation of planet's magnetic flux (planets are big magnets in space). In this way we pay for the energy induced on the wire by orbit decay. The orbit decay could be compensated by solar pressure if orbiting about the Sun.

   Figures 15 and 16 show two Flower Constellations architectures for power production in space. To my knowledge no constellation architecture has been developed or proposed for power production in space. In the current thinking, Space Solar Power Satellites require the launching and assembling in space of a very large structure in order to be economically viable. I believe that some Flower Constellation concepts can remove the need for in-space assembly. This architecture would replace the bulky approach to producing large areas where energy is collected into smaller and cheaper components. The advantage would be to provide a means of assembling a large collecting area without making it a grand challenge. Figures 15 and 16 show two view of a potential configuration. We outline about the possibility to re-orient the whole constellation and the possibility that each circle can be differently oriented. The dynamics is double: the circles spin about their centers and all of them spin (as rigid body) about the Flower Constellation axis of symmetry.

4. *Pointing Architectures* for a hyper large directional instrument through the use of Flower Constellations . These architectures would have the following functional capabilities:

   (a) use the possibility to align satellites using elliptical orbits (see Fig. 17 where more alignments are combined to form a more complex "star"-object), to promote directional active and passive observation and transmission of particles and energy on length

[4] Reference [11, 12] have shown that a GalileoSat GDOP-accuracy can be achieved by a GNFC made of 27 satellites, only. Recent investigations on designing GNFC using genetic algorithms provide even better results, bringing down to 26 (or even 25) the total number of satellites.

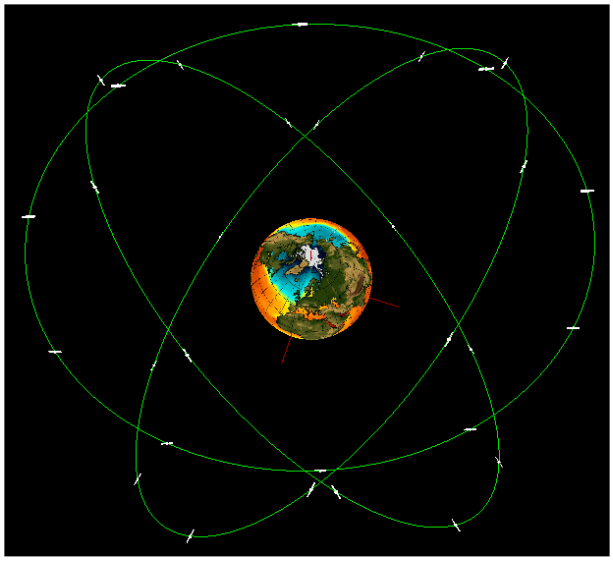


Figure 13: GalileoSat orbit planes

scale of the order of 1 AU or larger (passive observation can be performed through the pinpointing of a large lens assembly or the use of the interferometry configuration for the observation of planets of the solar system or other terrestrial planets outside the solar system)

(b) by collimating the Sun or other source of energy, one can provide active observation of materials of planets in our solar system by studying the properties of atmosphere and soil scattering of other planets with observatories on Earth. In particular, one could provide energy and particle transmission to future NASA mission probes beyond Jupiter by focusing energy or particles towards these probes, or provide a means of ablation to deflect asteroids.

(c) send signals beyond solar system by imitating common shapes on a very large scale.

(d) a small constellations fo surveillance and reconnaissance and well as for Space and Earth science.

5. *Laser Propulsion and Asteroid Deflection*. The idea is to design a Flower Constellation for laser propulsion of remote spacecraft using directional aligned lasers. The combination of lasers and satellites has captured the imagination of and interest of the scientific community. The potential applications, while futuristic at the current stage, are simply too important to discount: from laser beams that intercept and deflect earth-bound asteroids to laser propulsion in deep space, relaying and direction of the laser beams in the future will be accomplished by satellites. Despite the large advances in rocketry, the increases in payload capacity, and the effectiveness of orbit insertion, the possibility of putting massive lasers in orbit in the near future is minimal. If the laser beam pointing and steering can be accomplished by reflective and/or refractive elements on satellites, the potential exists for higher accuracy due to the lack of atmospheric beam steering, and reduction in the required laser power due to such accuracy. The Flower Constellation has one major advantage over other possible arrangements viz-a-viz laser beam manipulation: the existence of conserved paths with respect to the earth frame. One may think of ten satellites in a Flower Constellation as one large body in motion. Consequently, it is easier and more economical to aim one large-diameter beam towards the relaying Flower Constellation satellites than it is to aim several laser beams at every single satellite belonging in a general constellation. Severe beam decollimation occurs as laser beams propagate through the atmosphere, or even in vacuum. This is particularly true for the very powerful ultra-violet (UV) lasers that would have to be used for either laser propulsion or asteroid deflection. To some extent, the Flower Constellation turns this basic fact of laser propagation into an advantage: a circle of craft flying in a closed formation of a circle can be used for the relaying of a single laser beam of a given diameter. The precise choreography of the Flower Constellation craft makes it easier to generate these multiple laser beams in the first place.

The Flower Constellations represent a fundamental advance and a viable means to efficiently design new space objects, characterized by two, distinct, dynamics. The Flower Constellations represent a dramatic step forward with wide-ranging mission design impact, both

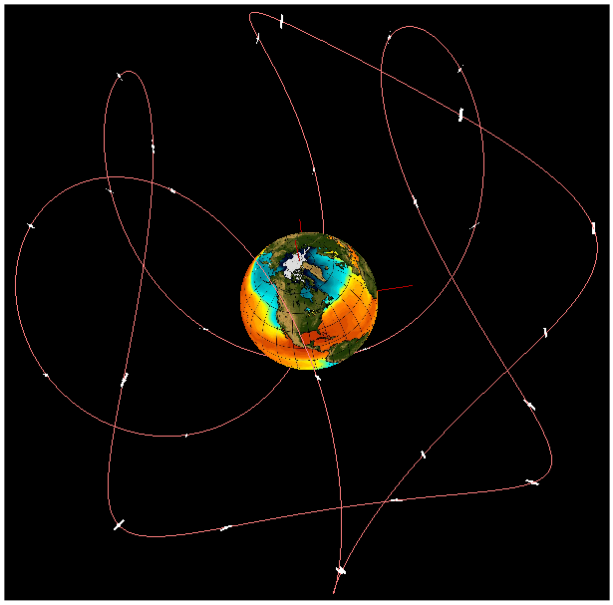


Figure 14: Flower Constellation Global Navigation

for future geocentric missions and the goals to move to the Moon, Mars, and beyond. In fact, new and more effective satellite constellations would strongly benefit many of the key strategic focus areas, already identified by NASA and ESA. In particular, we do expect to identify direct beneficiaries such as the Robotic and human lunar expeditions, the sustained, long-term robotic and human exploration of Mars, the robotic exploration across the solar system, the development of advanced telescopes searching for Earth-like planets and habitable environments, and the exploration of the Universe, of the dynamic Earth system, and of the Sun-Earth system.

It is obvious that in order to validate the proposed configurations, the analysis of the perturbations acting on each specific proposed constellations, must be done. Other validation criteria and performance metrics will be specific for each proposed solution. For instance, we may need to evaluate the type of particle/energy transmission capability that could benefit from these configurations (supercritical repeaters, collection of mirrors focalizing on one point) or evaluate the type of u-v-w plane capability for interferometry. In fact, in the interferometry systems currently evaluated by NASA, there is an expectation that a series of spacecrafts will be flying in constellation. This constellation flying constellation is likely to be restricted to a small baseline. We do expect to study how a line of sight between different spacecrafts can be attained using a Flower Constellation that enables a very large baseline. We do also expect to quantify the GNC system efforts needed to allow each spacecrafts to pinpoint in the right direction for interferometric observation.

Conclusion

Flower Constellations will have a large impact on future mission architectures and concepts. Flower Constellations are the 3-dimensional equivalents of "orbits with repeating ground tracks" which have been a staple for planet orbiting missions from Topex, EOS, to many of the planetary observation missions such as at Mars. Flower Constellations are ideal for studying 3-dimensional large scale structures and phenomena in space such as the detection of gravity waves, the study of magnetospheres, and radiation environment over vast regions around planets and moons. For example, such orbits could be used for the "Magnetospheric Constellation" that has been studied by NASA. Also, such orbits could provide the backbone for an "Inter-Planetary Network" for navigation and communication throughout the Earth-Moon system, at Mars and Jupiter, and eventually the entire Solar System. Such a "GPS" type network could enable the automation of many spacecraft functions including the adaptive on-board mission design and navigation in the future.

The 3-dimensional nature of these constellations enables the repeated visits of spatial locations to permit the study of the 3-dimensional structures of time varying phenomena in space around a planet. These constellations can be designed to enable complex, distributed instruments with virtual apertures extending 10s to 100s of km in diameter. As a bonus, if the axis of symmetry is aligned with the planet poles, these constellations actually have repeating ground tracks to boot.

The use of compatible orbits allows us to extend and enable the powerful techniques for studying planetary surfaces to 3-dimensional space around planets, as for

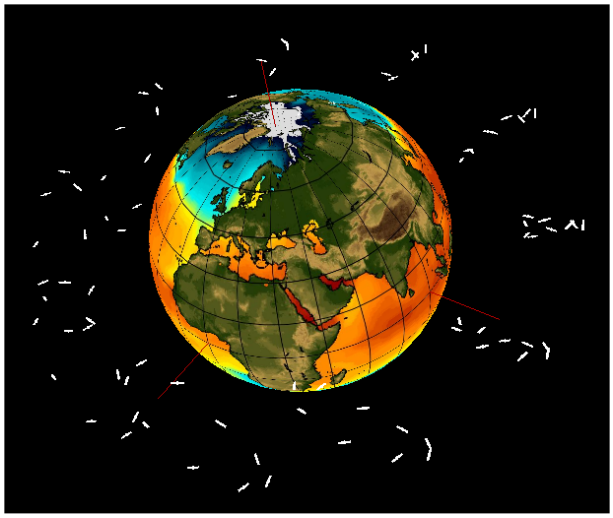


Figure 15: Flower Constellation Space Dynamo (view 1)

instance the planet's magnetosphere. By working with natural dynamics, Flower Constellations eliminate costly deterministic controls that limit current designs of more complex constellations around a planet.

Flower Constellations will have a dramatic impact on reducing cost and optimizing the functionality of satellite constellations for planetary exploration, as well as mapping the features of interest from orbit. The efficient use of natural dynamics reduces the number of control maneuvers required to maintain such constellations. This saves both propellant and operational costs. At the same time, the great variety of constellation patterns and 3 dimensionality of the constellation will enable new mission concepts and applications. Flower Constellations represent a fundamental advance in orbit design; the demonstration of a viable means to efficiently design Flower Constellations will represent a dramatic step forward with wide-ranging mission design impact, both for future geocentric missions and the goals to move to the Moon, Mars, and beyond.

Acknowledgments

The author would like to thank Dr. Matthew Paul Wilkins and Christian Bruccoleri for being the two leading columns in the Flower Constellations discovery and research. In particular, Dr. Wilkins, one of my former Ph.D. student, has been instrumental in the Flower Constellations discovering and in the theory development. The results of his efforts are now published on his Ph.D. dissertation [3]. To help the developing and the understanding of the Flower Constellation theory, my Ph.D. student Bruccoleri developed FCVAT [10], an ad-hoc software program to visualize and to design Flower Constellations . Without FCVAT it would be practically impossible to take the Flower Constellations at today's stage of maturity.

I want also to express my gratitude to other two of my Ph.D. students: Dr. Keun Joo Park, and Dr. Ossama Abdelkhalik, presently still working with me as a post-doc. In particular, Dr. Park has been the first to design a Flower Constellation for Global Navigation System (GNFC), while Dr. Abdelkhalik has applied the theory to ground reconnaissance problem from space and has developed the theory of the two-way orbits.

Bibliography

[1] Daniele Mortari, Matthew Paul Wilkins, and Christian Bruccoleri. The Flower Constellations. *Journal of the Astronautical Sciences*, 52(1 and 2):107–127, January–June 2004. Special Issue: The "John L. Junkins" Astrodynamics Symposium.

[2] Daniele Mortari and Matthew Paul Wilkins. Flower Constellation Set Theory Part I: Phasing and Compatibility. *IEEE Transactions on Aerospace and Electronic Systems*, 2006. Submitted for pubblication.

[3] Matthew Paul Wilkins. *The Flower Constellations – Theory, Design Process, and Applications*. PhD Thesis, Texas A&M University, October 2004.

[4] Troy Henderson and Daniele Mortari. Uni-Flower: A Novel Proposal for University-Built Nanosatellites in Flower Constellations. Paper AAS 06–204 of the 2006 Space Flight Mechanics Meeting Conference, Tampa, Florida, January 22–26 2006.

[5] Ossama Omar Abdelkhalik and Daniele Mortari. Two-Way Orbits. *Celestial Mechanics and Dynamic Astronomy*, 2006. To appear.

[6] Daniele Mortari, Ossama Omar Abdelkhalik, and Christian Bruccoleri. Synodic and Relative Flower

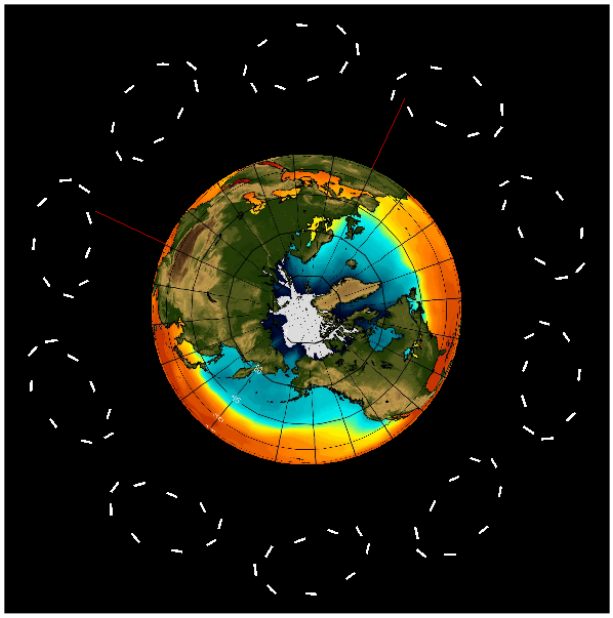


Figure 16: Flower Constellation Space Dynamo (view 2)

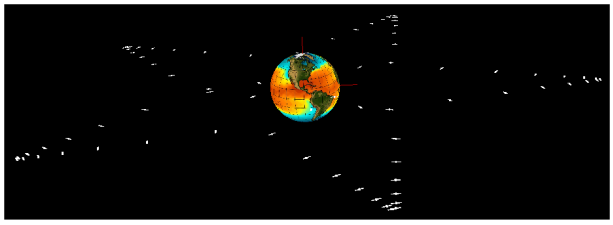


Figure 17: Lone Star Flower Constellation

Constellations with Applications for Planetary Explorations. Paper AAS 05-151 of the 2005 Space Flight Mechanics Meeting Conference, Copper Mountain, Colorado, January 23–27 2005.

[7] Daniele Mortari and Matthew Paul Wilkins. Dual-Compatible Flower Constellations. Paper AAS 06–202 of the 2006 Space Flight Mechanics Meeting Conference, Tampa, Florida, January 22–26 2006.

[8] Ossama Omar Abdelkhalik and Daniele Mortari. Orbit Design for Ground Surveillance Using Genetic Algorithms. *Journal of Guidance, Control, and Dynamics*, 2006. To appear.

[9] Matthew Paul Wilkins and Daniele Mortari. Flower Constellation Set Theory Part II: Secondary Paths and Equivalency. *IEEE Transactions on Aerospace and Electronic Systems*, 2006. Submitted for pubblication.

[10] Christian Bruccoleri and Daniele Mortari. The Flower Constellation Visualization and Analysis Tool. Paper of the 2005 IEEE Aerospace Conference, Big Sky, Montana, March 5–12 2005.

[11] Keun Joo Park, Matthew Paul Wilkins, Christian Bruccoleri, and Daniele Mortari. Uniformly Distributed Flower Constellation Design Study for Global Positioning System. Paper AAS 04-297 of the 2004 Space Flight Mechanics Meeting Conference, Maui, Hawaii, February 9–13, 2004.

[12] Keun Joo Park, Marina Ruggieri, and Daniele Mortari. Comparisons Between GalileoSat and Global Navigation Flower Constellations. Paper of the 2005 IEEE Aerospace Conference, Big Sky, Montana, March 5–12 2005.

[13] H. Le Coroller, J. Dejonghe, C. Arpesella, D. Vernet, and A. Labeyrie. Tests with a Carlina-type Hypertelescope Prototype: I. Demonstration of Star Tracking and Fringe Acquisition with a Balloon-Suspended Focal Camera. A&A, 2004.

Equilibrium Shaping

Dario Izzo[1] and Lorenzo Petazzi[2]

[1] Research Fellow, Advanced Concepts Team, European Space Agency, ESTEC, Keplerlaan 1, 2201 AZ Noordwijk, The Netherlands.
Email: dario.izzo@esa.int
[2] Ph.D candidate, ZARM, University of Bremen, Germany

Abstract

We present a satellite path planning technique able to make identical spacecrafts acquire a given configuration. The technique exploits a behaviour-based approach to achieve an autonomous and distributed control over the relative geometry making use of limited sensorial information. A desired velocity is defined for each satellite as a sum of different contributions coming from generic high level behaviours: forcing the final desired configuration the behaviours are further defined by an inverse dynamic calculation dubbed Equilibrium Shaping. We show how considering only three different kind of behaviours it is possible to acquire a number of interesting formations and we set down the theoretical framework to find the entire set. We find that allowing a limited amount of communication the technique may be used also to form complex lattice structures. Several control feedbacks able to track the desired velocities are introduced and discussed. Our results suggest that sliding mode control is particularly appropriate in connection with the developed technique.

Introduction

"Does a coherent group behaviour require an explicit mechanism of cooperation?" "Can useful tasks be accomplished by a homogeneous team of mobile agents without direct communication and using decentralized control?" These questions [1] have been addressed by an increasingly large community of computer scientists, engineers and scientists in general working in a field of research that we may call swarm intelligence or collective robotics. The relevance of the possible answers to the aerospace community is significant. Space engineers are currently developing autonomous systems and are envisaging space missions that would certainly benefit from a deeper understanding of the collective behaviour of similar and dissimilar agents. Multi-robot planetary exploration, on-orbit self-assembly and satellites swarm for coordinated observations are just examples of what could be achieved if our technology level was proved to be sufficient to provide spacecraft swarm with autonomous decision capabilities. As recently proposed by Ayre et al. [2] it would be possible to build large solar panels or large antennas exploiting collective emerging behaviours. The use of collective robotics is also very relevant for advanced missions architectures such as those recently studied by ESA (APIES mission [3]) and NASA (ANTS mission [4]) making use of satellite swarm to explore the asteroid belt.

Because of these synergies between collective robotics and space mission design, it makes sense to try the design of a decentralized control for a satellite swarm relying upon the lesson learned from collective robotics. Drawing inspiration from this research field we will often use the name agent to indicate a certain spacecraft belonging to some group. When a complex system of many agents has to act in a coordinated manner, the action selected by each component may or may not take into account the decisions taken by the others. The smallest the number of communications required between the agents, the smallest is the degree of coordination of the system. On the other hand a small amount of communications leads to a simple and robust system. The information exchanged with the other swarm components is useful but not necessary in defining the geometric and kinematical representation of the time varying environment that will then influence the agent action selection. Many works dealing with terrestrial robots navigation [5] with spacecraft proximity and rendezvous operation [6] and self-assembly structures in space [7] have taken the approach of defining an artificial potential field to model the environment. With this method the action selection

is then made by following the local gradient of the artificial potential field. Although this method allows a precise modelling of the external environment it also introduces local undesired equilibrium configurations that the system may reach. A laplacian-based potential field [8] or the use of harmonic functions [9] along with the introduction of random walks [10] have been proposed in order to alleviate the problem. Another approach to the action selection problem was introduced by Schoner, [11] based on the dynamic systems theory. In this approach the state space contains behavioural variables such as heading directions or velocities. All the contributions given by each behaviour are combined by means of weighting parameters into a final dynamical system that defines the course of behaviours that each agent will follow. The weighting parameters can be evaluated by solving a competitive dynamics operating at a faster time scale. Recently, also other approaches have been proposed for space applications in the attempt to obtain some degree of decentralized coordination in a group of satellites. Lawton and Beard [12, 13] introduced what they call a Virtual Structure method to design a decentralized formation control scheme while Campbell [14] applied some results of the optimal control theory in order to design a coordinated formation reconfiguration manoeuvre. These methods aim at reaching a unique final configuration in which each satellite has its position preassigned. When a swarm of homogeneous agents is considered and the task is given to acquire a certain final geometry, the final positions occupied by each agent in the target configurations should be chosen in an autonomous way and should be part of the global behaviour emerging from the individual tasks assigned.

In this paper we investigate the possibility of using the limited and local sensing capabilities of each single spacecraft to coordinate the individual responses and achieve a common task. The common task we try to achieve is the acquisition of a given relative geometry in which positions are not pre-assigned to particular satellites. We develop a behaviour based **path planning algorithm** [15] able to achieve this. Based on selective sensory information, each behaviour contributes to the final decision taken by the spacecraft control system. With respect to previously developed techniques the new approach presented has two advantages: it limits to a minimum amount inter-satellite communications and it autonomously assigns the final satellites positions.

Methodology

Consider the relative motion of N spacecraft randomly distributed in the space neighboring N targets and subject to the gravitational attraction of a near planet. In the Local Horizontal Local Vertical (LHLV) reference frame associated to a given target orbit we define the target positions $\boldsymbol{\xi}_i$, $i = 1..N$ and the initial states $\mathbf{x}_i^0, \mathbf{v}_i^0$, $i = 1..N$ of each spacecraft. Our goal is to build a real time navigation scheme allowing each agent to autonomously decide what final target to acquire relying just upon its limited sensor information, and to safely navigate to it without conflicting with the other spacecraft. We will follow a two-step approach:

- First, a method is developed that defines for each target disposition and each agent neighborhood configuration the desired velocity vector of the agent as a sum of different weighted contributions named "behaviours".
- Then, several control techniques are considered that allow each spacecraft to track the desired kinematical field.

In this way the control design is completely independent from the design of the desired velocity field and may be faced separately.

Design of the underlying kinematical field: the Equilibrium Shaping

The approach we propose and that we call Equilibrium Shaping, draws inspiration from past published works on robot path planning and artificial intelligence. In the work by Gazi [16, 17] some theoretical results have been introduced on the dynamics of aggregating swarm of robots. Each agent of the swarm is there asked to follow a certain velocity field defined as the sum of two different contributions, both solely dependent from the inter-agent distance $\mathbf{x}_{ij} = \mathbf{x}_j - \mathbf{x}_i$. The first contribution defines a linear global gather behaviour whereas the second one introduces an avoidance behaviour. The mathematical definition used by Gazi for the desired velocity of the $i-th$ agent is:

$$\mathbf{v}_{d_i} = -\sum_j \mathbf{x}_{ij} \left[c_i - b_i \exp\left(-\frac{\mathbf{x}_{ij} \cdot \mathbf{x}_{ij}}{k_1} \right) \right]$$

where c_i, b_i are coefficients whose values are uniquely determined by the formation geometry. This method produces a swarm in which each agent is preassigned to a particular place in the final formation. It is possible to achieve a global swarm behaviour that also solves the target assignment problem autonomously. This may be achieved defining the desired kinematical field according to the Equilibrium Shaping approach proposed in this paper. This technique consists in building a dynamical system that has as equilibrium points all the possible agents permutations in the final target formation. The system is then used as a definition for the desired velocities. It is essentially an inverse approach that starts from the geometry of the final relative configuration and works out the necessary agent behaviours to reach that configuration. Let us consider the simple example of a swarm of two satellites that aims at reaching a final configuration made up of the two geometric positions given by $\boldsymbol{\xi}_1 = [1, 0, 0]$ and $\boldsymbol{\xi}_2 = [-1, 0, 0]$. We have

to build a dynamical system that admits two equilibrium configurations, one in which the agent 1 is in $\boldsymbol{\xi}_1$ and agent 2 in $\boldsymbol{\xi}_2$ and one in which the final positions are inverted. We design our dynamic as a sum of three different behaviours that we name: "gather", "avoid"and "dock". The mathematical expression of each kind of behaviour along with some brief comments are listed below:

- **Gather Behaviour** This behaviour introduces N different global attractors towards the N targets. The analytical expression of this behaviour contribution to the i-th agent desired velocity may be written in the following form:

$$\mathbf{v}_i^{Gather} = \sum_j c_j \psi_G(\|\boldsymbol{\xi}_j - \mathbf{x}_i\|)(\boldsymbol{\xi}_j - \mathbf{x}_i)$$

  where ψ_G is a mapping from positive real to positive reals that introduces some non linear dependency from the target distance. This behaviour may also be designed to account for the gravitational field as we shall see in the next section. There are some important choices that we implicitly make when we choose this form of the gather behaviour. By allowing the c_j coefficient to depend solely on the targets and not on the agents we make sure that each component of the swarm is identical to the others so that agent permutations do not change the swarm behaviour. We also write the function ψ_G as dependant only on the distance so that an isotropy of the desired velocity field around each hole is imposed. This may not be desired in some particular problems in which case some angular dependency could be introduced.

- **Dock Behaviour** This behaviour introduces N different local attractors towards the N targets. The component of the desired velocity field due to each dock behaviour has a non-negligible value only if the agent is in the neighborhood of the sink. The k_D parameter determines the radius of the sphere of influence of the dock behaviour. The expression used for this behaviour is:

$$\mathbf{v}_i^{Dock} = \sum_j d_j \psi_D(\|\boldsymbol{\xi}_j - \mathbf{x}_i\|, k_D)(\boldsymbol{\xi}_j - \mathbf{x}_i)$$

  where ψ_G is a mapping from positive reals to positive reals that vanishes outside a given radius from the target. The same comments made for the gather behaviour still apply and the dock behaviour is similar to the gather one except that it is a local attractor and it therefore governs the final docking procedure.

- **Avoid Behaviour** This behaviour establishes a relationship between two different agents that are in proximity one with each other. In such a case a repulsive contribution will contribute to the desired velocity field. The expression that describes the desired velocity for this kind of behaviour is given below:

$$\mathbf{v}_i^{Avoid} = \sum_j b \psi_A(\|\mathbf{x}_i - \mathbf{x}_j\|, k_A)(\mathbf{x}_i - \mathbf{x}_j)$$

  where ψ_A is a mapping from positive real to positive reals that vanishes whenever the mutual distance is considered to be not dangerous according to the value k_A. In order to maintain the symmetry between all the agents the b parameter does not depend on the particular agent.

According to the definitions given before, the desired velocity field for a swarm of N agents and for a final formation made of N target is defined as follows:

$$\mathbf{v}_{d_i} = \mathbf{v}_i^{Avoid} + \mathbf{v}_i^{Dock} + \mathbf{v}_i^{Gather}. \tag{1}$$

This builds a dynamical system defined by the weighted sum of different and often conflicting behaviours and can be written in the simple form:

$$\dot{\mathbf{x}} = \mathbf{v}_{d_i} = \mathbf{f}(\mathbf{x}, \boldsymbol{\lambda}).$$

where we introduced $\mathbf{v}_d = [\mathbf{v}_{d_1}, .., \mathbf{v}_{d_N}]$, $\mathbf{x} = [\mathbf{x}_1, .., \mathbf{x}_N]$ and $\boldsymbol{\lambda} = [c_j, d_j, b]$. This last vector contains the parameters that have to be chosen so that all the final desired configurations are equilibrium points. As we took care of retaining the symmetry of the dynamical system with respect to agent permutations the only relation that has to be fulfilled in order to impose the existence of such equilibria can be written in the compact form:

$$\mathbf{f}(\mathbf{x}_e; \boldsymbol{\lambda}) = 0 \tag{2}$$

where $\mathbf{x}_e = [\xi_1, .., \xi_N]$. All the other configurations, obtained by permutation of the ξ_i, are granted to be also equilibrium points. This equation will be referred to as the Equilibrium Shaping formula, as it effectively allows to find the value of $\boldsymbol{\lambda}$ able to shape the equilibria of the dynamical system represented by Eq.(2). The study of what possible equilibria may be shaped with the previous equation reveals to be intriguing and well described by the theory of symmetry groups. Let us take a closer look at the Equilibrium Shaping formula. It is a set of N vectorial equations each one related to a particular target position ξ_i:

$$\sum_{\substack{j=1 \\ i \neq j}}^{N} \Big\{ \big[c_j \psi_G(\|\boldsymbol{\xi}_j - \boldsymbol{\xi}_i\|) + d_j \psi_D(\|\boldsymbol{\xi}_j - \boldsymbol{\xi}_i\|, k_D) - b \psi_A(\|\boldsymbol{\xi}_j - \boldsymbol{\xi}_i\|, k_A) \big] (\boldsymbol{\xi}_j - \boldsymbol{\xi}_i) \Big\} = 0 \tag{3}$$

It is convenient to think to b as a parameter, and to write the Equilibrium Shaping formula in the form:

$$\mathbf{A} \left[c_1, .., c_N, d_1, .., d_N \right]^T = \mathbf{g}$$

The matrix $\mathbf{A}$ and the vector $\mathbf{g}$ depend on the functions ψ chosen to represent the various behaviours and on the target positions. This set of equations represents, for each

target disposition and for each choice of the parameters k_D and k_A, a linear set of equations in the $2N$ unknowns c_j, d_j. Depending on the spatial distribution of the target points we might be able to find solutions. Let us further investigate the general case: a set of $3N$ equations in $2N$ unknowns and no possible solution. We must rely upon the linear dependency of some of the equations. We introduce the punctual symmetry group G of the target positions. Whenever two target points ξ_i, ξ_j are equivalent with respect to G (i.e. it exists a punctual symmetry belonging to G that maps ξ_i into ξ_j) then the two corresponding equations are linearly dependent if we set $c_i = c_j, d_i = d_j$. This statement can be easily proven pre-multiplying Eq.(3) by a matrix $R \in G$ and using the identity between the coefficients and the isometric property of R. This simple trick allow us to count the number of independent vectorial equations, each of them counting as three, two or just one scalar equation according to a simple rule. Each independent vectorial equation is equivalent to a single scalar equation whenever a symmetry axis passes through the considered point, as two scalar equations if a symmetry plan passes through it and it is an identity if more than one symmetry axis passes through it.

As an example let us consider the hexagonal Bravais lattice shown in Figure 1. The target positions belonging to this formation can be divided in two different groups (the prism vertices and the bases centers) belonging to two different symmetry classes. We therefore start with two linearly independent vectorial equations. As a symmetry plane passes through the vertices the corresponding vectorial equation counts as two equations. As a symmetry axis passes through the hexagons centers the corresponding vectorial equation counts as one equation and the Equilibrium Shaping formula may therefore reduced to a total of three scalar equations if we set $c_i = c_j, d_i = d_j$ among the positions belonging to the same group.

In Table 1 some formations are listed together with the number of independent equations and unknowns that characterize them. Some of the formations presented in Table 1 draw the inspiration from the well known Bravais lattice spatial configurations. The visualization of them is provided in Figure 2. Whenever the number of equations is less than the number of unknowns, the choice of the various parameters may be done in many different ways. Exploiting this fact the choice may be made differently by each agent and may be seen as the "subjective" view that the i-th spacecraft has of the equilibrium condition. There is no need to agree on the particular solution chosen.

Let us consider the case in which the target positions $\boldsymbol{\xi}_i$ form an icosahedron. For this particular target geometry (see Table 1 under regular solids) the Equilibrium Shaping formula reduces to a single scalar relation:

$$\begin{aligned} c = [5(R-h)e^{-\frac{\ell^2}{k_1}} + 5(R+h)e^{-\frac{\ell'^2}{k_1}} \\ + 2Re^{-\frac{2R^2}{k_1}}]b - [5(R-h)e^{-\frac{\ell^2}{k_2}} \\ + 5(R+h)e^{-\frac{\ell'^2}{k_2}} + 2Re^{-\frac{2R^2}{k_2}}]d \end{aligned} \quad (4)$$

where the quantities ℓ and R are respectively the edges of the icosahedron and the radius of the sphere in which the icosahedron can be inscribed, whereas ℓ' and h can be written as follows

$$\begin{aligned} h &= \sqrt{R - (\tfrac{\ell}{2\sin(\frac{\pi}{10})})^2} \\ \ell' &= \sqrt{(R+h)^2 + (\tfrac{\ell}{2\sin(\frac{\pi}{10})})^2}. \end{aligned}$$

For a fixed parameter b, Eq.(4) defines a relation between the two remaining parameters c and d. Any choice of these two parameters leads to a dynamical system describing the desired velocity having the icosahedron as equilibrium point. Such a dynamical system has the following form:

$$\begin{aligned} \dot{\mathbf{x}}_i &= \textstyle\sum_{j=1}^{N}[-b\exp(-\mathbf{x}_{ij}\cdot\mathbf{x}_{ij}/k_1)]\mathbf{x}_{ij} \\ &\quad + \textstyle\sum_{j=1}^{N}[-c - d\exp(-\boldsymbol{\xi}_{ij}\cdot\boldsymbol{\xi}_{ij}/k_2)]\boldsymbol{\xi}_{ij}. \end{aligned} \quad (5)$$

where we have defined $\boldsymbol{\xi}_{ij} = \boldsymbol{\xi}_i - \mathbf{x}_j$. The expression used for the function ψ is the one proposed by Gazi [16, 17]. It is possible to define ψ in different ways in order to decrease the computational load as done for example in [18] where a simple sine function is used. In Figure 3 the outcome of a numerical integration of this dynamical system is shown in the case $b = 1.1\frac{1}{s}$, $d = 0.3\frac{1}{s}$, $k_1 = k_2 = 1m^2$, $l = 2m$. The initial positions of the various agents have been randomly generated on the surface of a sphere of radius 20 times the length of the edge of the icosahedron. The lines shown in Figure 3 are the trajectories that each agent, having chosen d and c, foresees and uses to evaluate its desired velocity. We note that the agents do not perform any numerical integration to plan its path, only a simple algebraic calculation. During various simulations the rise of emerging behaviours due to the interaction between different conflicting behaviours [15] may be observed. As a final remark we note here that using the functions ψ proposed by Gazi [16] may lead to local minima configurations i.e. equilibrium configurations different form the desired one. A proper choice of the b, c and d parameters in the space of the solutions of the Equilibrium Shaping formula alleviates significantly this problem.

Exploiting the gravitational environment

The desired kinematical field designed in the previous section allows to reach the final formation following forced trajectories that are not geodesics. It is not difficult to imagine that the control system will struggle to follow these trajectories using unnecessarily large amount

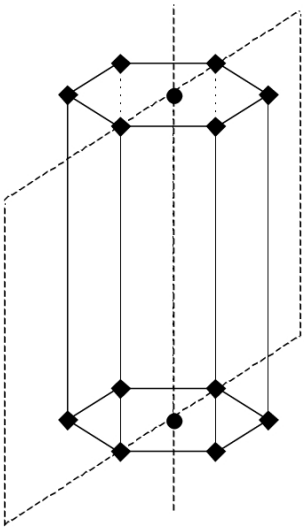


Figure 1: Formation with an Hexagonal-P shape.

Formation shape	Number of equations	Number of unknowns
All regular solids	1	2
All regular polygons	1	2
Pyramids with a regular basis	3	4
Cubic-P	1	2
Cubic-I	1	4
Cubic-F	2	4
Tetragonal-P	2	2
Tetragonal-I	2	4
Orthorhombic-P	3	2
Orthorhombic-I	3	4
Orthorhombic-C	4	4
Orthorhombic-F	6	4
Hexagonal-P	3	4

Table 1: Count of the equations and the unknowns for different formations.

of propellant whenever the gravitational forces become significant. A modification may be introduced that takes into account and exploits the geodesics to reduce the overall mass consumption. We start from the well-known system of Hill equations:

$$\begin{cases} \ddot{x} - 2\omega\dot{y} - 3\omega^2 x = 0 \\ \ddot{y} + 2\omega\dot{x} = 0 \\ \ddot{z} + \omega^2 z = 0 \end{cases}$$

which admit an exact analytical solution in the form:

$$\begin{bmatrix} \boldsymbol{\rho} \\ \dot{\boldsymbol{\rho}} \end{bmatrix} = \begin{bmatrix} \mathbf{A}(\tau) & \mathbf{B}(\tau) \\ \mathbf{C}(\tau) & \mathbf{D}(\tau) \end{bmatrix} \begin{bmatrix} \boldsymbol{\rho}_0 \\ \dot{\boldsymbol{\rho}}_0 \end{bmatrix}$$

where $[\boldsymbol{\rho}, \dot{\boldsymbol{\rho}}]$ is the non dimensional state space vector, τ the non dimensional time and the matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{D}$ have a well-known expression [7]. The above solution can be used in order to define a new gather behaviour that exploits the gravitational force to reach the final desired configuration. Requiring that a given satellite has to reach a certain point $\boldsymbol{\rho}_d$ after a fixed time τ_d the following relation has to be fulfilled:

$$\boldsymbol{\rho}_d = \boldsymbol{\rho}(\tau_d) = \mathbf{A}(\tau_d)\boldsymbol{\rho}_0 + \mathbf{B}(\tau_d)\dot{\boldsymbol{\rho}}_0.$$

Taking this into account we may assign for each position in the space $\mathbf{x}_i$ and each target belonging to the final formation $\boldsymbol{\xi}_j$ a new gather velocity vector given by:

$$\mathbf{v}_i^{Gather} = \frac{1}{N}\sum_j \mathbf{B}^{-1}\boldsymbol{\xi}_j - \mathbf{B}^{-1}\mathbf{A}(\hat{\tau}_d - t)\mathbf{x}_i \qquad (6)$$

where $\hat{\tau}_d$ is the time in which, at the beginning of the simulation, the agent is required to reach the center of the desired formation. Even though the resulting desired velocity vector depends explicitly on the time, in a practical application the agent clocks need not to

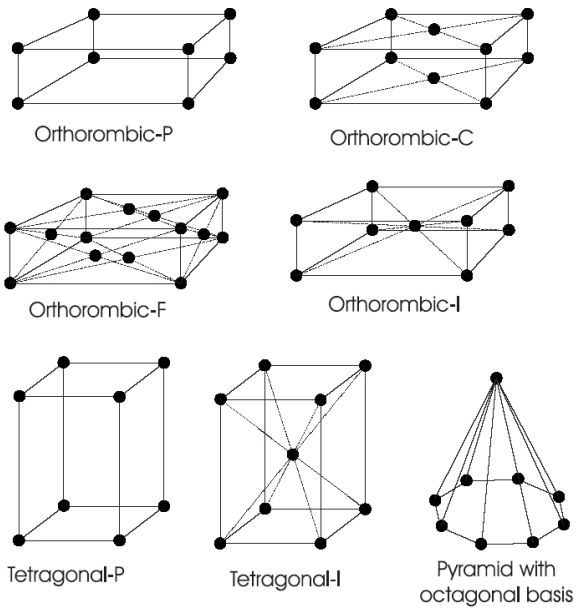


Figure 2: Visualization of some formations.

be synchronized. This contribution is added to the dock behaviour and the avoid behaviour in order to build the final desired kinematical field. Unfortunately Eq.(6) is singular when t approaches $\hat{\tau}_d$, i.e. in the final part of the target acquisition. Besides, near the targets the desired velocity due to this new gather behaviour is higher than needed (the spacecraft has to get out of a ballistic trajectory to acquire the targets). For these reasons the desired kinematical field will be divided in two different parts, one, far from the desired final configuration, in which the gather behaviour takes into account the gravitational force and one, close to the desired final formation, in which the space can be considered flat. The geometrical shape of the edge of these two different zones of the space can be easily set as a sphere of radius R_s that can be considered, together with the desired gather time $\hat{\tau}_d$ as a parameter to be decided by the system designer.

Feedback synthesis

In general the agent will not possess the desired velocity and a control system has to be designed that is able to reduce the error between the actual velocity and the desired one. In this section different feedbacks achieving this will be derived and discussed.

Q-guidance

The first feedback we develop is inspired by the Q-guidance steering law introduced formally by Battin [19] for rockets guidance. It is based on the definition of the "velocity to be gained"vector $\mathbf{v}_{g_i}$ that represents, in our case, the instantaneous difference between each agent's actual $\mathbf{v}_i$ and desired velocity $\mathbf{v}_{d_i}$. The objective of the control system is to drive the velocity to be gained vector to zero. From now on each quantity will be related to each agent but, in order to simplify the notation the subscript will be omitted. We define, for each agent, the following function:

$$V = \frac{1}{2}\mathbf{v}_g \cdot \mathbf{v}_g$$

the velocity to be gained vector decreases along the trajectories followed by each agent if and only if:

$$\dot{V} = \mathbf{v}_g \cdot \dot{\mathbf{v}}_g < 0. \tag{7}$$

The time derivative of $\mathbf{v}_g$ during the motion has the expression $\dot{\mathbf{v}}_g = \dot{\mathbf{v}}_d - \dot{\mathbf{v}}$. We substitute into this relation the momentum balance of each spacecraft written in the LHLV frame introduced:

$$\dot{\mathbf{v}} = \mathbf{f}_{in} + \mathbf{u}$$

where $\mathbf{f}_{in}$ are the external inertial forces and $\mathbf{u}$ is our control vector. The following expression is obtained:

$$\dot{\mathbf{v}}_g = \dot{\mathbf{v}}_d - \mathbf{f}_{in} - \mathbf{u}.$$

We now express the desired velocity derivative using the chain rule:

$$\dot{\mathbf{v}}_d = \frac{\partial \mathbf{v}_d}{\partial t} + \frac{\partial \mathbf{v}_d}{\partial \mathbf{x}}\mathbf{v} = \frac{\partial \mathbf{v}_d}{\partial t} + \frac{\partial \mathbf{v}_d}{\partial \mathbf{x}}(\mathbf{v}_d - \mathbf{v}_g)$$

The various terms can be found by deriving Eq.(1) and taking into account the expressions chosen for the different behaviours. As an example we consider the contribution

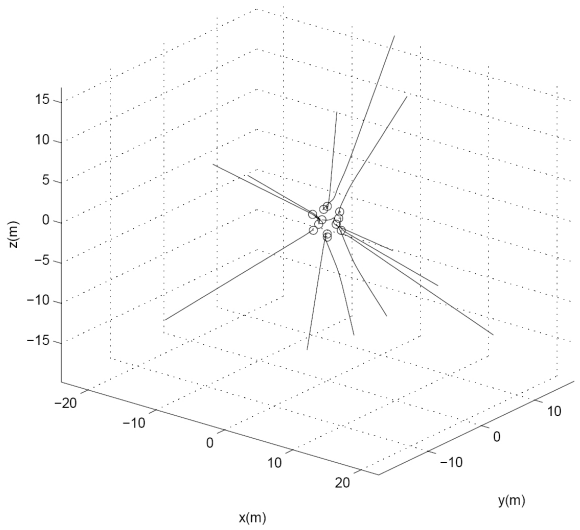


Figure 3: Example of desired trajectories for an icosahedron-shaped target formation.

given by the gather behaviour defined by Eq.(6). In this particular case the time derivative of the desired velocity has the expression:

$$\dot{\mathbf{v}}_d = \mathbf{f}_{in} - \frac{\partial \mathbf{v}_d}{\partial \mathbf{x}} \mathbf{v}_g$$

and the resulting total time derivative of the velocity to be gained (under the hypothesis that the avoid behaviour contribution is negligible) may be written in the following compact form:

$$\dot{\mathbf{v}}_g = -\mathbf{u} - \frac{\partial \mathbf{v}_d}{\partial \mathbf{x}} \mathbf{v}_g = -\mathbf{u} - \mathbf{B}^{-1}\mathbf{A}\mathbf{v}_g. \tag{8}$$

To make $\dot{V} < 0$ we introduce the following feedback:

$$\mathbf{u} = \kappa \mathbf{v}_g - \dot{\mathbf{v}}_d - \mathbf{f}_{in} \tag{9}$$

in which $\kappa > 0$ is a positive real parameter whose choice will be discussed below. We then have:

$$\dot{V} = -\kappa \mathbf{v}_g \cdot \mathbf{v}_g$$

that assures us the global stability of this controller. In real cases the agents will have an upper limit to the thrust magnitude, we may therefore have to saturate the feedback defined above. As soon as we introduce a saturation level we lose the mathematical result on the feedback global stability, but the controller is still able to drive the velocity to be gained to zero as confirmed by numerical simulations. In particular we here note how the geometrical interpretation of the saturated feedback shown in Figure 4 allows us to derive again, for particular choices of the positive parameter κ the analogous to known steering laws based on the Q-guidance (see Battin [19]).

If $\kappa \to \infty$ the control strategy is to thrust in the direction of the velocity to be gained vector regardless of the contribution to the $\dot{\mathbf{v}}_g$ due to the uncontrollable terms $\dot{\mathbf{v}}_d$ and $\mathbf{f}_{in}$. A different strategy can be achieved if the thrust direction is chosen in order to try aligning the time derivative of the velocity to be gained vector to the $\mathbf{v}_g$ vector itself, as expressed by the following relation:

$$\dot{\mathbf{v}}_g \times \mathbf{v}_g = 0. \tag{10}$$

We have reduced to the known cross product steering that, in our notation, may be implemented by finding that value of κ for which (see Figure 4):

$$(\dot{\mathbf{v}}_d + \kappa \mathbf{v}_g) \cdot (\dot{\mathbf{v}}_d + \kappa \mathbf{v}_g) = u_{sat}$$

where u_{sat} is the saturation considered for the thrust vector modulus.

Sliding mode control

A feedback can be also obtained using the results of the sliding-mode control theory (see for example [20]). The aim of this approach is to design a control law able to drive the system trajectory on a predetermined manifold and keep it there once reached. The control design procedure can be broken down in two different steps: first we design a sliding manifold (or switching manifold) such that the motion of the system restricted to it leads the swarm of satellites towards the desired equilibrium configuration. Then a control law has to be derived to force the trajectories of the system to collapse on the sliding manifold and to belong to it during the whole simulation.

The dynamical system to be controlled for each agent is

$$\begin{cases} \dot{\mathbf{v}} = \mathbf{f}_{in} + \mathbf{u} \\ \dot{\mathbf{x}} = \mathbf{v} \end{cases} \tag{11}$$

where $\mathbf{f}_{in}$ and $\mathbf{u}$ are respectively the inertial acceleration acting upon the agent and the control vector. The sliding

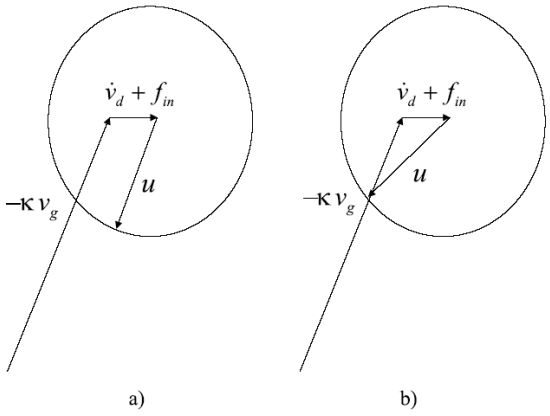


Figure 4: Vectorial diagrams representing different strategies.

manifold may be written in the following form:

$$\boldsymbol{\sigma}(\mathbf{x}, \mathbf{v}, t) = 0$$

where

$$\boldsymbol{\sigma}(\mathbf{x}, \mathbf{v}, t) = [\sigma_1(\mathbf{x}, \mathbf{v}, t), ..., \sigma_3(\mathbf{x}, \mathbf{v}, t)] = 0.$$

The relations $\sigma_i(\mathbf{x}, \mathbf{v}, t)$ have to be chosen such that the trajectory that the system follows "sliding" on the manifold reaches the final desired formation. The Equilibrium Shaping technique we introduced in this paper is in fact a method to build such a sliding manifold. For the system in Eq.(11) the expression we use for $\boldsymbol{\sigma}(\mathbf{x}, \mathbf{v}, t)$ is:

$$\boldsymbol{\sigma}(\mathbf{x}, \mathbf{v}, t) = \mathbf{v}_d - \mathbf{v} = 0 \tag{12}$$

where $\mathbf{v}_d$ is defined in Eq.(1). Whenever the system is on the sliding manifold it will stay on it if and only if the following relation is satisfied at each instant:

$$\dot{\boldsymbol{\sigma}}(\mathbf{x}, \mathbf{v}, t) = 0$$

that is, according to Eq.(11) and to Eq.(12):

$$\dot{\mathbf{v}}_d - \dot{\mathbf{v}} = \dot{\mathbf{v}}_d - \mathbf{f}_{in} - \mathbf{u} = 0.$$

It is then possible to define the equivalent control $\mathbf{u}_{eq}$ as a feedback that keeps the state of the system on the manifold for all the time instants:

$$\mathbf{u}_{eq} = -\mathbf{f}_{in} + \dot{\mathbf{v}}_d. \tag{13}$$

A particular case is when the gravitational gather behaviour is the only contribution to the desired velocity. Then the control force reduces to zero since on the sliding manifold we have:

$$\dot{\mathbf{v}}_d = \frac{\partial \mathbf{v}_d}{\partial t} + \frac{\partial \mathbf{v}_d}{\partial \mathbf{x}} \mathbf{v}_d = \mathbf{f}_{in}.$$

The dynamical system in Eq.(11) subject to the equivalent control will never leave the sliding manifold after having intersected it once. It is now necessary to add to the equivalent control another term that acts when $\boldsymbol{\sigma} \neq 0$ and is able to drive the system trajectory to intersect the sliding manifold. The total control vector can then be expressed as the sum of two contributions:

$$\mathbf{u} = \mathbf{u}_{eq} + \mathbf{u}_N \tag{14}$$

where $\mathbf{u}_{eq}$ has been defined in Eq.(13). The vector $\mathbf{u}$ applied to the system Eq.(11) couples the dynamics of each single spacecraft to the other components of the swarm that the spacecraft can sense. A switched control law in the form:

$$u_{N_i} = \begin{cases} u^+_{N_i}(\mathbf{x}, \mathbf{v}, t) \longleftrightarrow \sigma_i(\mathbf{x}, \mathbf{v}, t) > 0 \\ u^-_{N_i}(\mathbf{x}, \mathbf{v}, t) \longleftrightarrow \sigma_i(\mathbf{x}, \mathbf{v}, t) < 0 \end{cases}$$

will enforce the system to fall onto the sliding manifold if the values $u^+_{N_i}, u^-_{N_i}$ are chosen so that the velocities of the system will point always towards it. The value chosen for the feedback gains is determined according to the sign of the components of $\boldsymbol{\sigma}(\mathbf{x}, \mathbf{v}, t)$, also called switching surface for this reason. The control u_{N_i} is set to zero on the switching surface. A Lyapunov method can be used to find the values of the switching gains. Let's define the following Lyapunov function

$$V = \frac{1}{2} \boldsymbol{\sigma} \cdot \boldsymbol{\sigma}$$

then a control feedback must be derived thus to impose the time derivative of V to be negative definite along the trajectories of the system. The condition on the total time derivative of the Lyapunov function can be imposed by

$$\dot{V} = (\mathbf{v}_d - \mathbf{v}) \cdot (\dot{\mathbf{v}}_d - \dot{\mathbf{v}}) < 0$$

that, recalling Eq.(13) and Eq.(14), becomes

$$\dot{V} < 0 \Longleftrightarrow (\mathbf{v}_d - \mathbf{v}) \cdot \mathbf{u}_N > 0.$$

The latter equation can be written in terms of the velocity to be gained vector already defined

$$\mathbf{v}_g \cdot \mathbf{u}_N > 0. \tag{15}$$

Each additional feedback law $\mathbf{u}_N$ that meets this condition can be used in order to drive the motion of the system towards the sliding manifold. Consistently with the work presented by Gazi [16] and to keep the derivation close to the classical sliding mode approach, the so called "relays with constant gain" (see [20]) thrusting strategy for $\mathbf{u}_N$ is introduced:

$$\mathbf{u}_N = u_0 sign(\boldsymbol{\sigma}) = u_0 sign(\mathbf{v}_g) \tag{16}$$

where the $sign$ function is defined componentwise. This definition clearly satisfies Eq.(15) and leads the system to reach the sliding manifold and then the desired equilibrium configuration. As a final remark we note that the thrusting strategy $\mathbf{u} = \kappa \mathbf{v}_g + \dot{\mathbf{v}}_d - \mathbf{f}_{in}$ inspired by the Q-guidance method can be written as $\mathbf{u} = \mathbf{u}_{eq} + \mathbf{u}_N$ with $\kappa \mathbf{v}_g$ satisfying Eq.(15). In this sense the sliding mode theory and the velocity to be gained approach reveal to be equivalent.

Artificial Potential Approach

A different thrusting strategy can be obtained starting from the definition of an artificial potential function [6] for the whole swarm $V(\mathbf{x}_1, \cdots, \mathbf{x}_n, \mathbf{v}_1, \cdots, \mathbf{v}_n)$ that has minimum points in all the possible agents permutation in the final desired formation. Such a function of the state of the system can be written as:

$$V = \frac{1}{2}\sum_i \mathbf{v}_i \cdot \mathbf{v}_i + \sum_i \sum_{j \neq i} \phi_A^{ij}(\mathbf{x}_{ij}) + \sum_i \sum_j \phi_G^{ij}(\boldsymbol{\xi}_{ij}) + \sum_i \sum_j \phi_D^{ij}(\boldsymbol{\xi}_{ij}) \tag{17}$$

where ϕ_A^{ij}, ϕ_D^{ij} and ϕ_G^{ij} are defined according to the Equilibrium Shaping technique so that:

$$\begin{aligned} \frac{\partial \phi_A^{ij}}{\partial \mathbf{x}_i} &= -\mathbf{v}_i^{Avoid} \\ \frac{\partial \phi_G^{ij}}{\partial \mathbf{x}_i} &= -\mathbf{v}_i^{Gather} \\ \frac{\partial \phi_D^{ij}}{\partial \mathbf{x}_i} &= -\mathbf{v}_i^{Dock} \end{aligned}$$

and each quantity labeled with the i index is related to the $i-$th agent. The swarm will reach the target formation avoiding inter-vehicles collisions whenever the function $V(\mathbf{x}_1, \cdots, \mathbf{x}_n, \mathbf{v}_1, \cdots, \mathbf{v}_n)$ strictly decreases during the motion. We get:

$$\dot{V} = \sum_i \left(\frac{\partial V}{\partial \mathbf{x}_i}\dot{\mathbf{x}}_i + \frac{\partial V}{\partial \mathbf{v}_i}\dot{\mathbf{v}}_i\right) = \sum_i (\dot{\mathbf{v}}_i - \mathbf{v}_{d_i}) \cdot \mathbf{v}_i < 0.$$

Taking into account the $i-$th agent equation of motion it is possible to use the following feedback $\mathbf{u}_i = \mathbf{v}_{d_i} - \kappa_i \mathbf{v}_i - \mathbf{f}_{in_i}$ that written in terms of the velocity to be gained vector becomes:

$$\mathbf{u}_i = \kappa_i \mathbf{v}_{g_i} + (1 - \kappa_i)\mathbf{v}_{d_i} - \mathbf{f}_{in_i}. \tag{18}$$

With this feedback the time derivative of the potential function is:

$$\dot{V}(\mathbf{x}_1, \cdots, \mathbf{x}_n, \mathbf{v}_1, \cdots, \mathbf{v}_n) = -\sum_i \kappa_i \mathbf{v}_i \cdot \mathbf{v}_i$$

which is definite negative as long as the κ_i parameters are chosen positive. With respect to the previous presented feedback design methods the one showed in this subsection relies upon a slightly different approach. First a global artificial potential function is defined for the entire swarm of satellites. This function is required to have minimal positions in all the possible swarm target formations and this may be obtained using the Equilibrium Shaping approach. Then a control law is imposed such as the potential function decreases along the trajectories followed by the system. The feedback derived in this section cannot be obtained from the Q-guidance or the sliding mode and therefore represents an alternative to be considered.

Simulation Results

In this section we present some numerical simulations we performed to study the performances of the discussed behaviour based control. We randomly placed N satellites within a certain range and we activated the controller to study the swarm behaviour. We performed our simulations for different relative geometries. We also considered different feedbacks given by:

- Eq.(9) with κ such that Eq.(10) is satisfied. We called this feedback CPSL (Cross Product Steering Law)
- Eq.(14) with u_0 tuned in such a way as to make the acquisition time of the final targets equal to the CPSL case. We called this feedback SMC (Sliding Mode Control).
- Eq.(9) with κ tuned as for the SMC. We called this feedback VTBG (Velocity To Be Gained)
- Eq.(18) with κ tuned as for the SMC. We called this feedback APF (Artificial Potential Feedback).

Typically, as a consequence of the control actuation, each agent path consisted of different phases:

- A powered part in which the initial velocity to be gained vector is driven to zero and in which a ballistic trajectory is reached.
- A coasting phase in which the desired velocity and the actual velocity are identical and the control system does not use the actuators.
- A last phase, activated within the sphere of radius R_s, in which the agent is near to the final targets and navigates towards one of them and the final geometry is acquired. In this phase the gather behaviour does not take into account the gravitational effect.

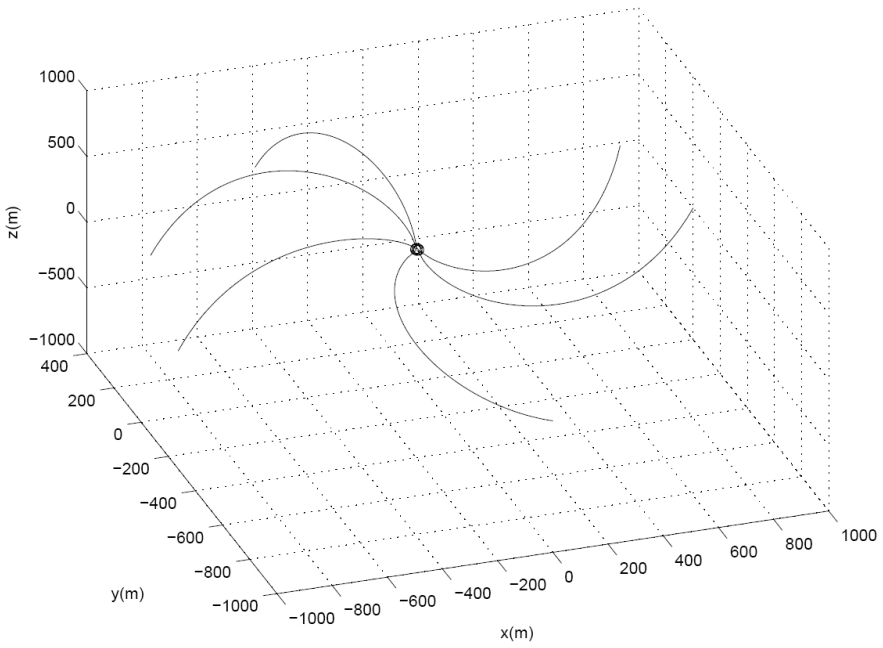


Figure 5: Simulation of a swarm of satellites reaching an hexagonal regular formation. First approaching phase

The final relative geometry is achieved only when every agent occupies a target position so that the Equilibrium Shaping formula is satisfied and the desired velocities are all zero. As an example we show the trajectories followed by six satellites achieving an hexagonal formation with a $6m$ radius. Each spacecraft belonging to the swarm starts from an average distance of $1000m$ with respect to the center of the final configuration. The saturation value used for the thrust acceleration modulus is $0.005\frac{m}{s^2}$ and the feedback law used is the VTBG. The final formation was achieved after roughly $20000sec$ corresponding to roughly a quarter of the reference orbital period. In Figure 5 and 6 the trajectories followed by the spacecraft belonging to the swarm are displayed. Figure 5 shows the motion of the swarm in the outer part of the kinematical field where gravity is accounted for, while Figure 6 shows the motion of the swarm in the very last phase. In this particular simulation the center of the desired formation was on a geostationary orbit. The thrust profiles of each spacecraft are shown in Figure 7. The different phases we described at the beginning of this section are visible in this chart. The expensive phases of the whole procedure in terms of propellant consumption are at the very first seconds when the engines are constantly saturated in order to reach a ballistic trajectory and at the last part of the formation acquisition when the gravitational force is no more considered in the definition of the desired velocities. It is in this phase that each agent chooses its final position and navigates towards it. An average Δv of about $0.8\frac{m}{s}$ was required in this particular manoeuvre by each agent.

In Figure 8 the different feedbacks introduced are considered and compared in terms of propellant consumption for this particular simulation. The numerical campaign performed showed that the SMC and the VTBG feedback are always outperforming the CPSL and the APF feedbacks.

Towards a complex lattice

In order to plan and control their path according to the Equilibrium Shaping approach each satellite, at least in a limited neighborhood, is required to detect the position of the neighboring spacecrafts. This sensing capability may in principle be provided in different ways according to the particular mission considered. Exploration mission [3] would probably use different solutions with respect to LEO formation flying missions or GEO self-assembly concepts [2]. Looking at future technological developments it seems plausible that wireless communication [21] will be an attractive option to be used on board missions implementing such a path planning scheme. Clearly once inter satellite communication is envisaged we may think to use it also for other purposes. In particular we here show how the use of a limited inter satellite communication may be used to overcome the limitations in the possible final geometries achievable by the equilibrium shaping approach. Let us consider a swarm of satellites divided into seeds and and non seeds agents. Each seed is associated to a number of non seed agents. Moreover let us introduce a new avoid behaviour that will be named as inner avoid and that will be effective at a smaller length scale i.e.

$$\mathbf{v}_i^{Avoid_{in}} = \sum_j b_{in}\psi_A(\|\mathbf{x}_i - \mathbf{x}_j\|, k_{A_{in}})(\mathbf{x}_i - \mathbf{x}_j)$$

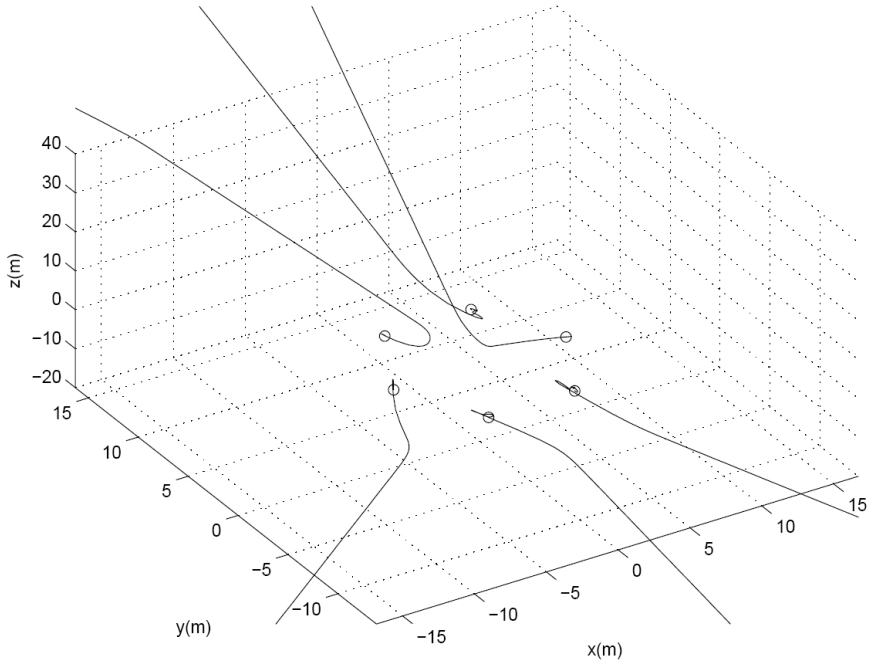


Figure 6: Simulation of a swarm of satellites reaching an hexagonal regular formation. Second approaching phase

with $k_{A_{in}} < k_A$. Then each spacecraft belonging to the swarm can evaluate its own desired velocity field according to the expression

$$\mathbf{v}_{d_i} = \mathbf{v}_i^{Avoid} + \mathbf{v}_i^{Avoid_{in}} + \mathbf{v}_i^{Dock} + \mathbf{v}_i^{Gather}.$$

The seeds plan their path ignoring the non seed agents (except for the inner collision avoidance). They evaluate and track their desired velocities $\mathbf{v}_{s_i}$. They do need to communicate their desired velocity to the other non-seed agents belonging to their group. These evaluate their desired velocities $\mathbf{v}_{a_j}$ using the Equilibrium Shaping approach but in a frame attached to the seed so that they will eventually form their formation around the seed. The control system of these non seed agents will try to track a velocity given by the sum of the two contributions: $\mathbf{v}_{tot_j} = \mathbf{v}_{s_i} + \mathbf{v}_{a_j}$. In this scheme each of the non seed agents need to communicate with the seed assigned to them to receive its actual desired velocity. Playing with the possible formation achievable by the Equilibrium Shaping it is in this way possible to build a scheme able to form complex lattices. The inner avoid behaviour ensures that agents belonging to different groups will not collide. However in order to have a system for which the Equilibrium Shaping formula can still be applied, the inner avoid behaviour must act at a length scale smaller than a characteristic length associated to the final configuration. With simple geometric considerations it is possible to find a value for the $k_{A_{in}}$ parameter such that there is no interaction between agents due to the inner avoidance behaviour when each spacecraft is in one of the final positions of the target configuration. Under this condition the Equilibrium Shaping formula can be applied as it appears in eq.(3).

In the following the principal features of the already described path planning scheme are shown through two different examples. To simplify the simulations each agent is supposed to have perfect control i.e. it can perfectly track the desired kinematical field and the desired kinematical field does not take into account the gravity field. We take as a first example a group of 72 agents divided into 8 groups of 8 non seed agents and a group of 8 agents considered as seed. In this case, we performed simulations placing the agents randomly in the proximity of a large target configuration made by a cubic lattice. In the simulation shown the average distance of the agents from the center of the desired formation at $t = 0$ was $477.7m$ whereas the standard deviation was $137.8m$. The satellites are able to plan their path and acquire the final desired configuration in which the seeds are disposed to form a cube of side $L = \frac{1}{\sqrt{3}} \cdot 80m$ whereas each group of non-seed spacecraft forms a cube of side $l = \frac{1}{\sqrt{3}} \cdot 16m$ around each agent. The minimum distance between two agents belonging to the formation is $L - l$ so that the inner avoid behaviour must satisfy the condition $k_{A_{in}} < \frac{L-l}{2}$. For the cube formation the Equilibrium Shaping formula reduces to a sole independent equation in three unknowns. It is then possible to choose independently the values of the parameters b and d in the space of the solutions of the Equilibrium Shaping formula in order to select the formation acquisition time. When such parameters are set to be respectively $b = 0.055\frac{1}{s}$, $d = 0.0015\frac{1}{s}$ and

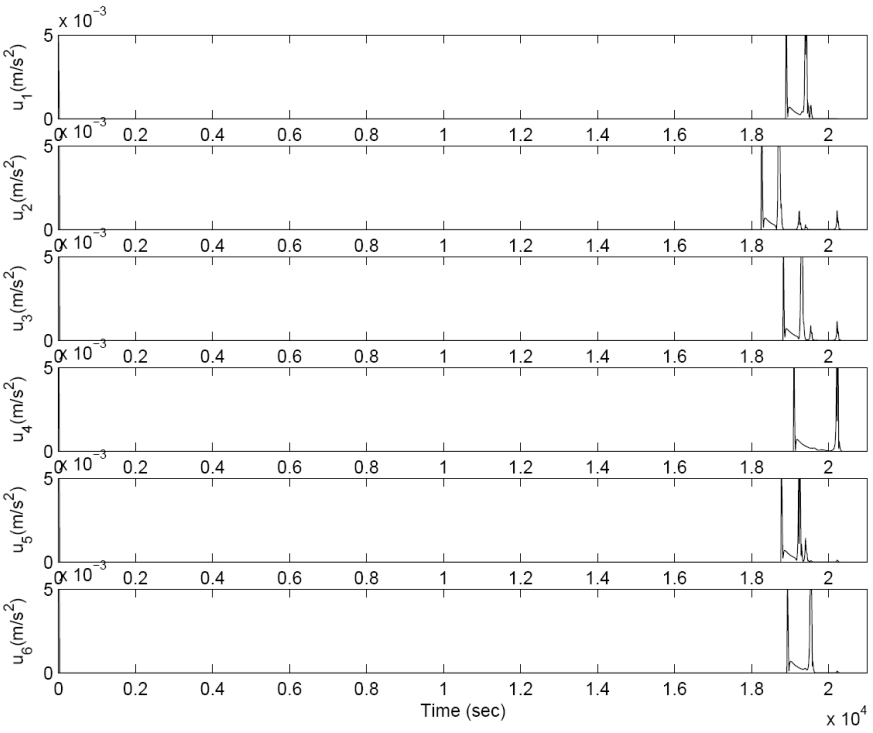


Figure 7: Control profiles for the satellites performing the maneuver.

$c = 1.899 \cdot 10^{-4} \frac{1}{s}$ for the seeds and $b = 0.0429 \frac{1}{s}$, $d = 0.00117 \frac{1}{s}$ and $c = 1.484 \cdot 10^{-4} \frac{1}{s}$ for the agents, the final formation is acquired in approximately $10000 sec$. In Figure 9 the actual assembly sequence is shown as returned by one of the performed simulations.

In the second example we integrate the already described method in a procedure aimed at driving the assembly of large structures in space. Let us consider for example a swarm of 49 spacecraft, with a group of 7 homogeneous seeds and 7 groups of 6 non seeds agents. In this example the target formation is a flat structure that can be used as a large reflector for deep space observation or solar power collection purposes [2]. In this case the reflector assembly can be accomplished by exploiting the application of the Equilibrium Shaping technique together with an autonomous docking algorithm. Figure 9 displays the different phases of the assembly procedure described in detail in the following. At the initial time the 49 agents are distributed in a cloud with an average distance with respect to the swarm center of mass of $453.8m$ and with a standard deviation of $147.1m$. An equilibrium shaping procedure brings the spacecraft belonging to the swarm to acquire a lattice formation in which the seeds are placed to form a centered hexagon of side $R = 20m$ while the non seeds agents are disposed around each seed as to form an hexagon of side $r = 4m$. The minimum distance between two agents belonging to different groups is $R-2r$ resulting in a condition for the $k_{A_{in}}$ parameter of $k_{A_{in}} < \frac{R-2r}{2}$. With respect to the previous simulation the formation of the seeds yields an Equilibrium Shaping formula that is made of a sole independent equation in five unknowns namely the d and c parameters for the target position at the vertices of the hexagon (marked as d_v and c_v), the ones for the position at the center of the hexagon (marked as d_c and c_c) and the b parameter for the avoidance behaviour. In particular the parameters d_c, c_c d_v and b have been chosen in order to minimize the occurrence of local minima. The Equilibrium Shaping formula written for the formation of the non seeds agents is made of a sole independent equation in three unknowns. The values used in this simulation are for the seeds $b = 0.06875 \frac{1}{s}$, $c_c = 0.02345 \cdot 10^{-3} \frac{1}{s}$, $d_c = 0.01875 \frac{1}{s}$, $d_v = 0.05 \frac{1}{s}$ and $c_v = 0.13162 \cdot 10^{-3} \frac{1}{s}$ whereas for the non seed agents $b = 0.04296 \frac{1}{s}$, $c = 8.3554 \cdot 10^{-4} \frac{1}{s}$ and $d = 0.01562 \frac{1}{s}$. With these parameters the swarm acquires the desired formation in approximately $20000 sec$. This phase of the assembly procedure is shown in figure 9 from frame 7 to frame 8. Once the final formation is acquired the spacecraft can perform the docking maneuver relying upon an autonomous docking procedure (figure 9 from frame 1 to frame 6). A detailed discussion on this procedure is out of the scope of the present paper. As a final remark we point out that the position of each agent in the final formation is not preassigned and it is autonomously decided during the simulation by the agents in a fully decentralized manner. In particular, in the simulations presented in this section, the swarm composed by N seeds and N groups of n agents autonomously decides according to the given initial conditions among the $(N!) \cdot (n!)^N$ possibilities which configuration to acquire in the space. The problem of self-assembling an hexagonal lattice has recently been solved also at a molecular level by using an inverse method

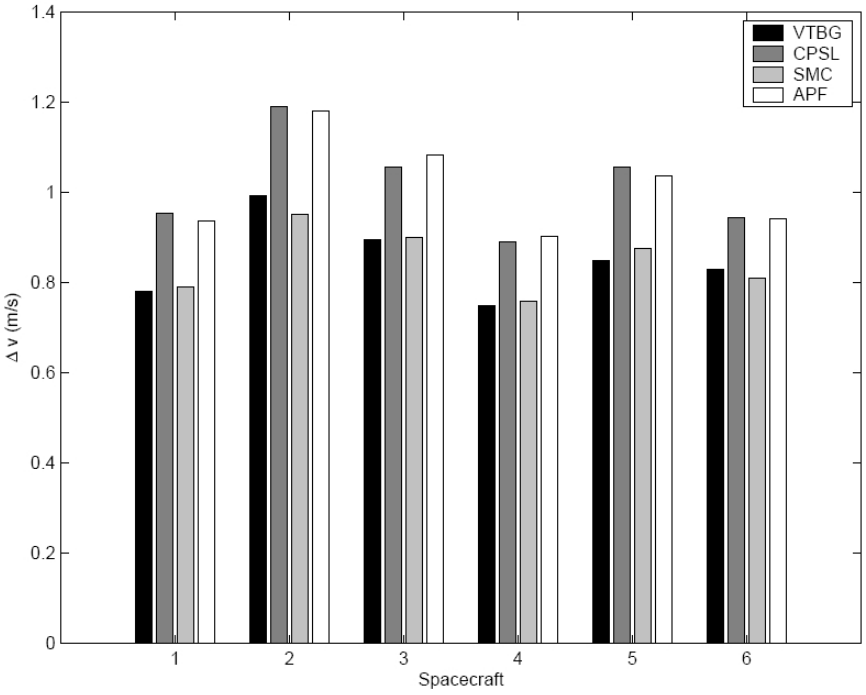


Figure 8: Comparison between the different thrusting strategies proposed.

by Rechtsman et al. [22]. Their technique is though not applicable to satellites navigation and path planning as it is inherently two dimensional. The similarity with the equilibrium shaping is anyway important and proves that inverse methods for self-assembly are interesting in connection with different fields.

Conclusions

We show that the task of acquiring certain formations may be achieved by a satellite swarm using only local sensory information. At the same time also the target selection problem may be solved so that the final position occupied by a given satellite does not have to be pre assigned. Using a behaviour based approach it is found that the introduction of three simple behaviours that have been named gather, avoid and dock, allow to reach a number of interesting formation geometries. These are found to be described properly by the symmetry group theory applied to the solutions of the Equilibrium Shaping formula here introduced. Some simulations of various control feedback show that the requirements in terms of propellant for real applications are well within our technological capabilities. More complicated structures may also be acquired at the cost of adding some extra inter satellite communication.

Acknowledgments

The authors wish to acknowledge the contributions given by Mark Ayre in developing the ideas included in this paper.

Bibliography

[1] C.R. Kube and H Zhang. Collective robotic intelligence. In *Proceedings of the Second International Conference on Simulation of Adaptive Behavior*, pages 460–468, 1992.

[2] M. Ayre, D. Izzo, and L. Pettazzi. Self assembly in space using behaviour based intelligent components. In *Proceedings of TAROS (Towards Autonomous Robotic Systems)*, Imperial College, London, UK, 2005.

[3] EADS-Astrium. The apies mission. Feasibility Study A0/1-3846/02/NL/JA, ESA, 2004.

[4] S.A. Curtis, J. Mica, J. Nuth, G. Marr, M. Rilee, and M. Bath. Ants (autonomous nano-technology swarm): An artificial intelligence approach to asteroid belt resource exploration. International Astronautical Federation, 2000. 51st Congress.

[5] Omar Khatib. Real-time obstacle avoidance for manipulators and mobile robots. *The International Journal of Robotics Research*, 5(1):90–98, 1986.

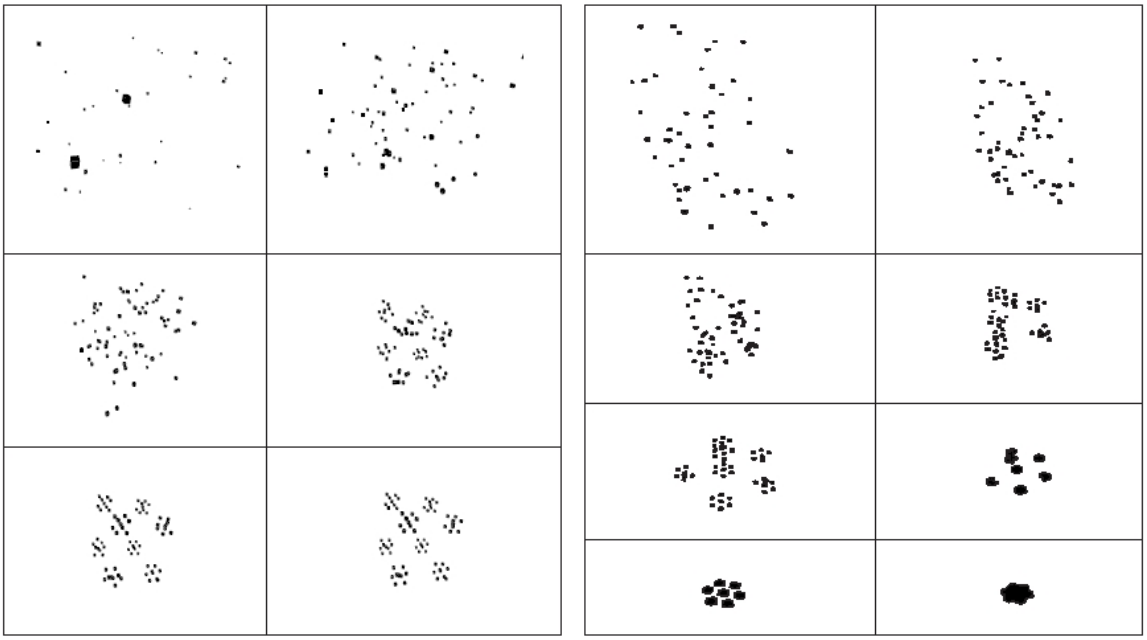


Figure 9: Assembly sequence for two lattices.

[6] Colin McInnes. Autonomous rendezvous using artificial potential functions. *Journal of Guidance Control and Dynamics*, 18(2):237–241, 1995.

[7] Frank McQuade. *Autonomous Control for On-Orbit Assembly Using Artificial Potential Functions*. PhD thesis, Faculty of Engineering University of Glasgow, 1997.

[8] K Sato. Dead-lock motion path planning using the laplace potential field. *Advanced Robotics*, 17(5):449–461, 1993.

[9] D Keymeulen and J. Decuyper. The fluid dynamics applied to mobile robot motion: the stream field method. In *Proceedings of the IEEE Int. Conference on Robotics and Automation*, pages 378–386, 1994.

[10] H. Chang. A new technique to handle local minimum for imperfect potential field based path planning. In *Proceedings of the IEEE Int. Conference on Robotics and Automation*, volume 1, pages 108–112, 1996.

[11] Gregor Schoner and M. Dose. A dynamics systems approach to task level systems integration used to plan and control autonomous vehicle motion. *Robotics and Autonomous Systems*, 10:253–267, 1992.

[12] W. Ren and R.W. Beard. A decentralized scheme for spacecraft formation flying via the virtual structure approach. *AIAA Journal of Guidance, Control and Dynamics*, 27(1):73–82, January 2004.

[13] J. Lawton, B. Young, and R. Beard. A decentralized approach to elementary formation maneuvers. *IEEE Transactions on Robotics and Automation*, 17(6):933–941, 2003.

[14] Mark Campbell. Planning algorithm for multiple satellite clusters. *Journal of Guidance Control and Dynamics*, 26(5):770–780, 2003.

[15] P. Pirjanian. Behavior coordination mechanism state of the art. Tech-Report IRIS-99-375, Institute for Robotics and Intelligent Systems, School of Engineering, University of Southern California, October 1999.

[16] Veysel Gazi. Swarm aggregations using artificial potentials and sliding mode control. In *Proceedings of the 42nd IEEE Conference on Decision and Control*, pages 2531–2536, December 2003.

[17] Veysel Gazi and Kevin Passino. A class of attraction/repulsion functions for stable swarm aggregations. *International Journal of Control*, 77(18):1567–1579, December 2004.

[18] Edward Large, Henrik Christensen, and Ruzena Bajcsy. Dynamic robot planning: Cooperation through competition. In *IEEE International Conference on Robotics and Automation*, volume 3, pages 2306–2312, 1997.

[19] Richard H. Battin. *An Introduction to the Mathematics and Methods of Astrodynamics*. AIAA Educational Series, New York, 1987.

[20] R.A. De Carlo, S.H. Zak, and S.V. Drakunov. *The Control Handbook*, chapter 57.5. CRC Press/IEEE Press, 1996.

[21] R. Magness. Short-range rf wireless (proximity) networks technology assessment for space application. Technology Dossier TOSE-1B-DOS-4 TOS-EDD/2004.3/RM, ESA, 2006.

[22] M.C. Rechtsman, F.H. Stillinger, and S. Torquato. Optimised interactions for targeted self-assembly: Application to a honeycomb lattice. *Physical Review Letters*, 95(228301), 25th November 2005.

Coulomb Force Virtual Space Structures

Gordon Parker[1], Hanspeter Schaub[2], Arun Natarajan[2] and Lyon King[1]

[1] Michigan Technological University, MEEM Dept., Houghton, MI 49931, USA
[2] Virginia Tech, Aerospace and Ocean Engineering Dept., Blacksburg, VA 24061, USA

Abstract

Recently, several uses for intercraft Coulomb forces have been explored. Proposed applications have ranged from creating static formations of many spacecraft to steerable nanosat deployment systems. This paper considers the use of Coulomb forces for creating space structures. Unlike conventional space structures, these "virtual space structures" have no physical connections. Instead, they are held together by maintaining specific charges at their node points. This form of structure is thus readily expandable, can be reconfigured to different shapes, and is easily deployed. Fundamental concepts are covered in conjunction with a control law illustrating the ability to form a structure with conventional stiffness and damping coefficients that can be actively modified.

Introduction

The forces generated between two charged bodies in space are sufficiently large that they can be used for a variety of useful purposes [1, 2]. A simplistic example is shown in Figure 1 where the two bodies are perfect spheres. In a vacuum the Coulomb force, f_{12}, is

$$f_{12} = \frac{k_c q_1 q_2}{d^2} \tag{1}$$

where k_c is Coulomb's constant ($8.99 \times 10^9 \frac{Nm^2}{C^2}$), q_1 and q_2 are the charges of the two bodies in Coulombs and d is the distance between their centers in meters.
Charged bodies in a plasma, such as in the vicinity of Earth, have a shielding effect characterized by the Debye length (λ_d). When a node actively charges, the oppositely charged particles in the plasma are attracted to it. As seen by other participating Coulomb force nodes, it is a cloud of particles with a net zero charge, and thus provides no Coulomb force. This Debye shielding phenomenon is modeled as an exponential decrease in Coulomb force with increasing separation distance. The ideal intercraft Coulomb force of Eq. 1 becomes

$$f_{12} = \frac{k_c q_1 q_2}{d^2} e^{-d/L_\lambda} \tag{2}$$

when considering the plasma Debye length shielding. For separation distances of $d > 2\lambda_d$ the Coulomb force effect is negligible.

The Debye length is a function of altitude, but is not a constant quantity due, for example, to changes in solar activity. In general λ_d is quite small at low Earth altitudes (centimeter level). At geostationary altitudes (GEO) it ranges from 10s of meters to 100s of meters. Thus, Coulomb force exploitation for creating virtual structures, or formation flying, is limited to either higher altitude operations, or interplanetary operations where again λ_d is large.

Active control of spacecraft charging has been an important area of study for many years. Differential charging is typically an undesirable phenomenon and can result in arcing between components and electronic failures. The node charging envisioned for a Coulomb virtual structure is not differential, but whole-craft charging. This is a far safer scenario than differential charging, where kilovolt level potentials should be possible.

The space plasma environment has both electrons and H^+ ions. Since the electrons are roughly 2000 times less massive than the ions, a body moving through the near-Earth plasma accumulates more electrons than ions, and thus, tends to charge negative. The active node charging for a Coulomb virtual structure will require kilovolt level charging. Although this seems prohibitively large, the SCATHA spacecraft, launched in 1979, demonstrated net "natural" potentials as high as -14kV as described by Mullen *et. al.*[3]. Using

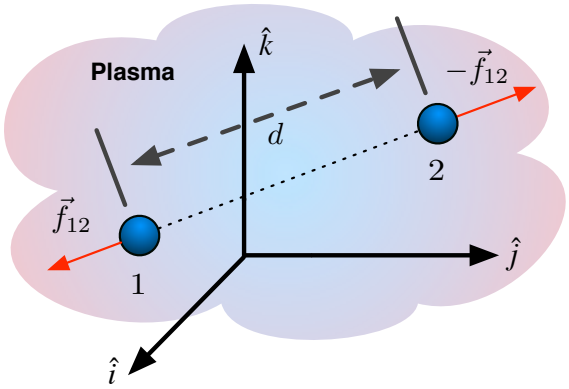


Figure 1: Coulomb forces between two charged spheres.

a low power electron emitter, SCATHA was actively charged to 3kV. Recent analysis shows that kilovolt level potential swings can be achieved with very little power and essentially zero propellant [1, 4]. Achieving additional negative charge, on demand, will require active ion emission. This can be accomplished with existing technology which has proven suitable for spacecraft operations [5].

The focus of this paper is the use of Coulomb forces for creating space structures. Since these structures do not have physical connections and are held together by the Coulomb forces alone, they will be called "virtual structures". A survey of recent activity in this area is presented. In addition, the analogy to conventional structures is illustrated with a simple two-node structure. Possible avenues for future research directions are also provided.

Virtual Structure Concept

A Coulomb virtual structure is comprised of several nodes whose charges can be actively controlled. Some nodes may have their own propulsion system for providing a net force to the structure's center of mass, while others may have instrumentation providing the scientific functionality of the structure. A typical structure is shown in Figure 2.

The modularity of this type of structure is apparent. Sections could be readily added, as long as the highly coupled charge node interaction effects are considered. Assuming that this system of N charged bodies is orbiting the Earth such that its center of mass has a nominal circular orbit, the Clohessy-Wiltshire [6] or Hill's equations [7] can be used to approximate their Hill-frame relative dynamics. These are shown in Eq. 3 for the N body case.

$$\begin{aligned} m\left(\ddot{x}_i - 2n\dot{y}_i - 3n^2 x_i\right) &= k_c \sum_{j=1}^{N} \frac{x_i - x_j}{d_{ij}^3} q_i q_j e^{-\frac{d_{ij}}{\lambda_d}} \\ m\left(\ddot{y}_i + 2n\dot{x}_i\right) &= k_c \sum_{j=1}^{N} \frac{y_i - y_j}{d_{ij}^3} q_i q_j e^{-\frac{d_{ij}}{\lambda_d}} \\ m\left(\ddot{z}_i + n^2 z_i\right) &= k_c \sum_{j=1}^{N} \frac{z_i - z_j}{d_{ij}^3} q_i q_j e^{-\frac{d_{ij}}{\lambda_d}} \end{aligned} \tag{3}$$

for $i = 1 \ldots N$, $j \neq i$ during the summation and $d_{ij} = \|\vec{p}_i - \vec{p}_j\|$. The Hill frame angular velocity is denoted as n and the mass of each node is m. The position of the ith node, relative to the Hill frame origin at the structure's center of mass, is denoted by the components of $\vec{p}_i$ given by x_i, y_i, z_i. Clearly, each charged node interacts with all the other charged nodes in a complex, but well defined way. Simply adding a new section to the structure without taking this into account would result in structural deformation, and eventually the structure's demise due to the orbital dynamic effects on the left side of Eq. 3. It should be noted, however, that for particularly large structures with characteristic lengths greater than $2\lambda_d$, distant nodes will no longer interact due to the Debye shielding effect. Thus, the control problem of actively modulating the node charges to achieve equilibrium could be simplified due to the Debye shielding effect.

Static Virtual Structures

One of the earliest attempts at exploiting Coulomb forces was for creating static, crystal-like spacecraft formations. At the heart of this analysis is the solution to Eq. 3 where the speed and acceleration terms are set to zero. The resulting static equilibrium equations are shown in Eq. 4

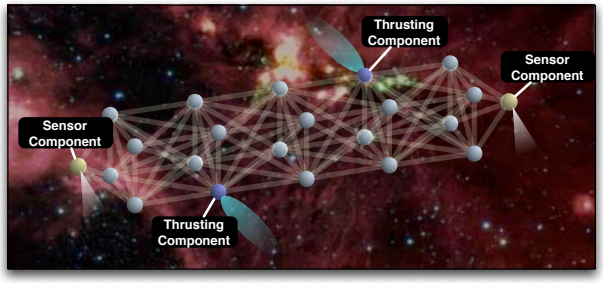


Figure 2: A Coulomb virtual structure illustrating the use of nodes not only for shape control, but also for net structure thrusting and scientific sensors.

$$
\begin{aligned}
-3\tilde{x}_i &= \sum_{j=1}^{N} \frac{\tilde{x}_i - \tilde{x}_j}{d_{ij}^3} \tilde{V}_i \tilde{V}_j e^{-M d_{ij}} \\
0 &= \sum_{j=1}^{N} \frac{\tilde{y}_i - \tilde{y}_j}{d_{ij}^3} \tilde{V}_i \tilde{V}_j e^{-M d_{ij}} \\
\tilde{z}_i &= \sum_{j=1}^{N} \frac{\tilde{z}_i - \tilde{z}_j}{d_{ij}^3} \tilde{V}_i \tilde{V}_j e^{-M d_{ij}}
\end{aligned}
\tag{4}
$$

where M is the number of Debye lengths considered, and both the positions and charges have been normalized using

$$
\begin{aligned}
\tilde{x}_i &= \frac{x_i}{M\lambda_d} \\
\tilde{y}_i &= \frac{y_i}{M\lambda_d} \\
\tilde{z}_i &= \frac{z_i}{M\lambda_d}
\end{aligned}
\tag{5}
$$

and

$$
\tilde{V}_i = \frac{r_{s/c}}{n\sqrt{m\left(M\lambda_d\right)^3 k_c}} V_i \tag{6}
$$

Chong[8] showed that analytical solutions to the static Coulomb equations of Eq. 4 exist. By enforcing specific shape symmetries she demonstrated structures with up to 6 participating nodes. It should be noted that these were not "free-flying" formations. Each had a central Coulomb node with its own conventional propulsion system. Thus, all the other nodes could react against the central node. Stability was considered after finding equilibrium configurations. It was concluded that none of the shapes was passively stable. Thus they would require an active control strategy to maintain their shape.

This original work illustrated that virtual structures of 10s of meters could be created using Coulomb forces. The next investigation focused on free-flying formations with none of the nodes having a conventional propulsion system described by Berryman and Schaub[9, 10]. Given a specified number of nodes, a genetic algorithm optimization approach was used to solve for the relative positions and charges such that static equilibrium equations were satisfied. Although there was no prescription as to the desired shape, this illustrated that free-flying virtual structures were possible with up to 9 nodes as shown in Figure 3. Larger clusters are certainly possible, using the optimization approach applied to the 9 node case. It should be noted that Debye length shielding was not considered. Necessary conditions on the shapes were also developed for satisfaction of the equilibrium equations [11]. In short, it was required that the structures principle inertia axes be collinear with the Hill frame.

This work was later extended to generating virtual structures that achieved a desired shape using a subset of the nodes as described by Parker *et. al.*[12]. In general, a Coulomb virtual structure will need more nodes than required of a conventional structure. These extra nodes are needed to generate the proper internodal forces to balance the static equilibrium equations of Eq. 4. For example, an eight node square box structure will not provide enough control degrees of freedom to satisfy the equilibrium equations.

Again, an optimization strategy was used to generate both the node charges and the positions of all the nodes. The specified shape nodes were incorporated as a penalty function in the overall cost function shown in Eq. 7 where $\vec{\tilde{R}}$ are the residual accelerations derived from Eq. 4. The first term of the cost function has two purposes. It simultaneously favors enforcement of the static equilibrium equations while trying to maintain even charges across all the nodes. The second term, with weighting factor w_1, enforces the desired shape function embodied in S. The last term favors satisfaction of the static formation necessary conditions. Debye length shielding was considered in this study where it was shown that specified shapes could be readily constructed. A specific example is shown in Figure 4 where the goal was to maintain three of the five nodes in a triangle as viewed by an observer on Earth located directly below the structure.

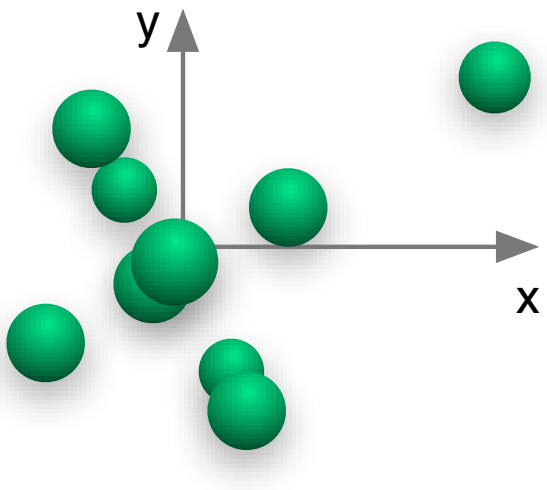


Figure 3: A nine node free-flying virtual structure.

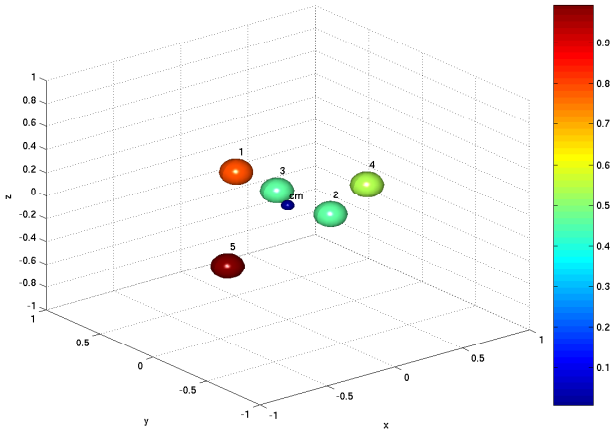


Figure 4: A five node structure with 3 nodes forming an equilateral triangle.

$$J = \left(\frac{\max \left| \tilde{V}_i \right|}{\min \left| \tilde{V}_i \right|} \right) \sum_{i=1}^{N} \left| \vec{\tilde{R}}_i \right| + w_1 S(\vec{\tilde{p}}_i, \vec{L_p}) + w_2 \Big\{ |I_{xy}| + |I_{xz}| + |I_{yz}| \Big\} \tag{7}$$

The color bar illustrates the normalized node voltages, scaled to that of the maximum node voltage. For an equilateral side length of 12 meters, a λ_d of 50 meters, and a node radius of 0.5 meters, the maximum node voltage was 16 kilovolts.

Virtual Structure Analogy to Conventional Space Structures

From the previous analyses two things are evident. First, practically sized virtual structures are feasible. Second, some form of closed-loop control will be needed to maintain the structure's shape. The difficulty of implementing a control strategy, due to the dynamic coupling between nodes, has already been mentioned. A charge cycling approach has been suggested for effectively decoupling the node interaction to all but two participating nodes at any one time [13]. In addition, the ability to create (and modify) conventional structure vibration characteristics in Coulomb virtual structure using closed-loop control is possible using the method described by Natarajan and Schaub[14].

Consider a simple, two node structure such that each node is nominally on the Hill frame x axis, that is, along a line directed radially out from the Earth. Taking as degrees-of-freedom the separation distance between the nodes (L), the in plane rotation of the structure (ψ) and the out of plane rotation (θ), the Hill dynamic equations can be represented as shown in Eq. 8 where δ denotes a small perturbation from the nominal configuration and $L_{\rm ref}$ is the desired separation distance that can, in general, be a time-varying quantity.

$$\begin{aligned} &\ddot{\theta} + \frac{2\dot{L}_{\rm ref}}{L_{\rm ref}}\dot{\theta} + 4n^2\theta = 0 \\ &\ddot{\psi} + \frac{2\dot{L}_{\rm ref}}{L_{\rm ref}}\dot{\psi} + \frac{2n}{L_{\rm ref}}\delta\dot{L} - \frac{2\dot{L}_{\rm ref}}{L_{\rm ref}^2}n\delta L + \frac{2\dot{L}_{\rm ref}}{L_{\rm ref}}n + 3n^2\psi = 0 \\ &\delta\ddot{L} + \ddot{L}_{\rm ref} - 2nL_{\rm ref}\dot{\psi} - 9n^2\delta L - \frac{k_c}{m_1}\delta Q\frac{1}{L_{\rm ref}^2}\frac{m_1+m_2}{m_2} = 0 \end{aligned} \tag{8}$$

The resulting set of linearized equations is time-varying if the reference separation distance, $L_{\rm ref}$ is specified to change. Postulating the control law of Eq. 9

$$\delta Q = \frac{m_1 m_2 L_{\rm ref}^2}{(m_1+m_2)\,k_c}(-C_1\delta L - C_2\delta\dot{L} + 2nL_{\rm ref}\dot{\psi}) \tag{9}$$

where the constants C_1 and C_2 are the position and velocity feedback gains, results in a closed-loop virtual structure which has a specifiable axial stiffness and damping. Specifically, the closed-loop δL equation becomes

$$\delta\ddot{L} + C_2\delta\dot{L}\left(C_1 - 9n^2\right)\delta L = 0 \tag{10}$$

The structure's axial mode natural frequency and damping ratio are

$$\begin{aligned} \omega_n &= \sqrt{C_1 - 9n^2} \\ \zeta &= \frac{C_2}{2\sqrt{C_1 - 9n^2}} \end{aligned} \tag{11}$$

and can be prescribed independently by selecting the control gains C_1 and C_2 appropriately.

This analogy is only meaningful if the entire system of Eq. 8 is stable. If $L_{\rm ref}$ is constant, then the stability bounds on C_1 and C_2 are readily computed by analyzing the characteristic equation of Eq. 8 and shows that $C_1 > 9n^2$ and $C_2 > 0$ are required. Stability of the time-varying system requires bounds on $\dot{L}_{\rm ref}$ and has been considered by Natarajan and Schaub[14].

Future Research Directions

In general the exploitation of Coulomb forces for spacecraft missions is an open area of research. Fundamental and experimental work is needed in the area of charge sensing and modulation. This would benefit the full spectrum of applications, including Coulomb virtual structures. Control strategies are needed to accommodate the complexity associated with the coupled nonlinear dynamics of an N node Coulomb virtual structure. A direct approach would be to devise a coupled nonlinear strategy that yields desired positioning performance while guaranteeing stability. Indirect methods may also prove effective where subsets of the structure are decoupled. The idea being that the solution to several reduced order problems would likely be more tractable than designing a control system for one high order system.

Active reconfiguration is a potential benefit of Coulomb virtual structures that has not been fully investigated. While conceptually easy, this may prove to be quite challenging. It is likely that exploitation of orbital dynamics will by beneficial. This may require development of trajectory design strategies that guarantee the controllability of the structure throughout its metamorphosis between shapes.

Consideration of the analogous structural properties of a Coulomb virtual structure may facilitate the use of existing space structure control techniques. This would require significant work beyond that described above for the simple two-node system. If successful, it may be possible to tailor the structure's natural frequencies and mode shapes to suit changing rigidity requirements. For example, while vibration sensitive scientific instruments

are active, its host subsystem may be decoupled from the rest of the structure, achieving true isolation.

Bibliography

[1] Lyon B. King, Gordon G. Parker, Satwik Deshmukh, and Jer-Hong Chong. Spacecraft formation-flying using inter-vehicle coulomb forces. Technical report, NASA/NIAC, January 2002. `http://www.niac.usra.edu`.

[2] Lyon B. King, Gordon G. Parker, Satwik Deshmukh, and Jer-Hong Chong. Study of interspacecraft coulomb forces and implications for formation flying. *AIAA Journal of Propulsion and Power*, 19(3):497–505, May–June 2003.

[3] E. G. Mullen, M. S. Gussenhoven, and D. A Hardy. Scatha survey of high-voltage spacecraft charging in sunlight. *Journal of the Geophysical Sciences*, 91:1074–1090, 1986.

[4] Hanspeter Schaub, Gordon G. Parker, and Lyon B. King. Challenges and prospect of coulomb formations. In *AAS John L. Junkins Astrodynamics Symposium*, College Station, TX, May 23-24 2003. Paper No. AAS-03-278.

[5] W. Riedler and et al. Active spacecraft potential control. *Space Science Reviews*, 79:271–302, 1997.

[6] W. H. Clohessy and R. S. Wiltshire. Terminal guidance system for satellite rendezvous. *Journal of the Aerospace Sciences*, 27(9):653–658, Sept. 1960.

[7] George William Hill. Researches in the lunar theory. *American Journal of Mathematics*, 1(1):5–26, 1878.

[8] Jer-Hong Chong. Dynamic behavior of spacecraft formation flying using coulomb forces. Master's thesis, Michigan Technological University, May 2002.

[9] John Berryman and Hanspeter Schaub. Static equilibrium configurations in geo coulomb spacecraft formations. In *AAS/AIAA Space Flight Mechanics Meeting*, Copper Mountain, Colorado, Jan. 2005. Paper No. AAS 05-104.

[10] J. Berryman and H. Schaub. Analytical charge analysis for 2- and 3-craft coulomb formations. Lake Tahoe, Aug. 2005. Paper No. 05-278.

[11] Hanspeter Schaub, Christopher Hall, and John Berryman. Necessary conditions for circularly-restricted static coulomb formations. In *AAS Malcolm D. Shuster Astronatuics Symposium*, Buffalo, NY, June 12-15 2005. Paper No. AAS-05-472.

[12] Gordon Parker, Lyon King, and Hanspeter Schaub. Charge determination for specified shape coulomb force virtual structures. In *47th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, Newport, RI, May 1-4 2006.

[13] Gordon Parker, Lyon King, and Hanspeter Schaub. Steered spacecraft deployment using interspacecraft coulomb forces. In *American Control Conference*, Minneapolis, MN, June 14-16 2006.

[14] Arun Natarajan and Hanspeter Schaub. Linear dynamics and stability analysis of a coulomb tether formation. In *15^{th} AAS/AIAA Space Flight Mechanics Meeting*, Copper Mountain, Colorado, Jan. 2005. Paper No. AAS 05-204.

Interstellar Precursor Missions using Advanced Dual-Stage Ion Propulsion Systems

Dave G. Fearn[1] and Roger Walker[2]

[1]EP Solutions, 23 Bowenhurst Road, Church Crookham, Fleet, Hants, GU52 6HS, UK. EMail: dg.fearn@virgin.net
[2]Advanced Concepts Team, ESA/ESTEC, Keplerlaan 1, 2201 AZ Noordwijk, The Netherlands. EMail: Roger.Walker@esa.int

Abstract

In this paper it is shown that an advanced form of gridded ion thruster, employing a novel 4-grid ion extraction and acceleration system rather than the usual two or three grid variants, can provide a velocity increment and specific impulse of interest to interstellar precursor missions, extending to a few hundred astronomical units from the sun. In this it is assumed that a nuclear power source is available with a mass-to-power ratio of 15 to 35 kg/kW and an output of at least several tens of kW. Mission durations are of about 25 years and the velocity increment provided exceeds 37 km/s. The paper includes a description of the technical approach adopted to achieving the required values of specific impulse, thrust density and power consumption, and presents for the first time the data obtained from an experimental programme conducted at ESTEC to verify the principles on which these theoretical predictions are based.

1. Introduction

An important parameter in assessing whether an interstellar precursor mission, or one to the outermost planets, is viable in programmatic terms is the time required for the receipt by the ground segment of data of scientific interest. It is normally assumed that this time must be well within the career span of a senior scientist working on the project at the outset; he must be confident that his early endeavours will be rewarded by the acquisition of such data before he retires. Similarly, those responsible for funding the mission will have comparable requirements. This translates into a mission duration of not more than about 25 to 30 years.

Since the velocity increment, ΔV, necessary to perform such a mission is very large, high spacecraft velocities are mandatory. It should be noted in this context that gravity assist manoeuvres to supplement on-board propulsion are not recommended in general terms, because of severely restricted launch windows. It can thus be concluded that either a very energetic launch is required, with an attendant high cost, or that the spacecraft must carry its own effective propulsion system, which can provide the necessary velocity increment. In the latter case, launch costs can be reduced further by deploying the spacecraft into a low altitude Earth orbit, then using the on-board propulsion system to increase the orbital radius until escape velocity is achieved1[1].

A simple application of Newton's Laws of Motion shows that the only way in which to avoid a very large propellant mass is to operate at an extremely high specific impulse (SI), implying the use of electric thrusters. This paper assesses the requirements of such missions, and concludes that the only near-term candidate propulsion technology which might be applicable is the gridded ion engine, which is very flexible, in that the exhaust velocity, and thus the SI, can be varied at will. However, a power source is also required, and it must be assumed that this will be nuclear, since the performance of solar arrays at such distances from the sun is not acceptable. A modified Topaz-2 type of reactor[2], utilising turbomachinery for power conversion, can in principle provide several tens of kW to the ion propulsion system, with a mass to power ratio of the order of 22 to 27 kg/kW.

With these assumptions, an optimisation process can be carried out for any mission of interest to determine the value of SI required to minimise the launch mass. This procedure is illustrated and typical parameters are reviewed. It is shown how the necessary ion engine design can be achieved, using a 4-grid technology which

permits the ion extraction and acceleration processes to be separated and therefore controlled independently. Values of SI reaching as high as 150,000 s can be attained, should that be required, and thrust levels can reach the Newton range. The principles involved in designing such an ion engine have been evaluated experimentally very recently at ESTEC, with most encouraging results. These have shown that ion extraction and acceleration to energies as high as 30 keV can be realised with a simple laboratory test arrangement. In these initial experiments, thrust densities of the order of 9 mN/cm^2 have been achieved, together with values of SI approaching 18,000s when using xenon propellant, and ion beam divergences of 2°. All three values represent world best performances for gridded ion engines.

2. The Gridded Ion Thruster

2.1 Principles

A schematic diagram of a gridded ion thruster is shown in Fig 1. It consists essentially of a discharge chamber, in which the propellant gas is ionised, and a set of ion extraction and acceleration grids. The energy required for the ionisation process is provided via a direct current (DC) or a radiofrequency (RF) discharge. A triple-grid system is shown, because this configuration is superior to the alternative twin-grid design as regards resistance to damage from bombardment by charge-exchange ions[3]; it therefore provides a significantly longer lifetime and causes less sputtered material to be deposited upon spacecraft surfaces. The positive space-charge of the ion beam is neutralised by electrons extracted from an external plasma created in a DC discharge between a neutraliser hollow cathode and an adjacent keeper electrode. The ion beam current, and thus the thrust, is controlled by the rate at which propellant is supplied and by the ionisation process, and the exhaust velocity, and hence the SI, is controlled by the potential of the beam power supply. The middle (accel) grid is at a negative potential of a few hundred volts relative to the spacecraft potential to prevent external electrons from back-streaming into the thruster.

The essential power supplies are indicated in Fig 1. The beam and accel grid supplies are used to extract the ions from the discharge plasma and then accelerate them to high velocity. As mentioned above, a third (decel) grid can be added to maximise lifetime. This can utilise another power supply, with a negative output, or can be connected to the spacecraft electrical ground. The beam supplies are voltage regulated and must be sufficiently robust to withstand shorts caused by occasional inter-grid arcs, although they are rare under space vacuum conditions. The neutraliser requires a heater supply, to raise the temperature of the cathode prior to discharge initiation[4], and a keeper supply to achieve breakdown and then to maintain the subsequent discharge.

A major feature is the separation of the three main processes required to produce a high energy ion beam. The first is the plasma production process, the second is the extraction and acceleration of the positive ions, and the third is the neutralisation of the space-charge of the ion beam. This feature has very significant design and operational advantages. For example, different grid sets can be designed to provide a wide variety of thrusts and values of SI, utilising the same discharge chamber and neutraliser system. Then, in operation, a very wide throttling range[5] is available, merely by altering the discharge chamber parameters.

2.2 Operation at Very High SI and Power

In all present thrusters the ion extraction system is of either twin- or triple-grid configuration; the latter is depicted in Fig. 1. Such grids operate in an accelerate/decelerate mode to enhance the throughput of ions. In this scheme, the thruster body and screen grid are at the high potential, V_B, necessary to achieve the ion beam velocity required. For example, if xenon ions are to be emitted at 40 km/s (SI $\sim$ 3500 s), a potential of 1.1 kV is needed. The accel grid is at a negative potential, of perhaps - 250 V, to focus the beamlets, to enable the required current to be extracted, and to prevent electron back-streaming from the external plasma. Deceleration to space potential then follows; with triple-grids, this is via a less negative voltage applied to the decel grid, which is typically - 50 V.

To raise the ion velocity, ν_e, and thus the SI, it is merely necessary to increase V_B as appropriate. However, this causes a greater penetration of the inter-grid electric field into the discharge chamber plasma. Some penetration is desirable, since this results in a curved plasma sheath which both increases the ion emission area and aids focusing. Unfortunately, the curvature becomes severe at high voltages, influencing the ion trajectories adversely and causing direct impingement on the outer grids. This situation is depicted in Fig 2, in which the sheath positions for a moderate and a high electric field, E, are shown. Direct ion impingement on the accel grid is evident in the latter case. Moreover, the plasma number density, n, in the discharge chamber must also be taken into account, since the penetration of the sheath increases as n falls. This problem is relevant to the use of a wide throttling range, or if the radial plasma density distribution is strongly peaked. However, in the latter case the dimensions of the screen grid apertures can be matched to this distribution to alleviate adverse effects[6]. Thus it is clear that high values of E and/or low values of n allow the sheath to penetrate deeply into the plasma. The resulting curvature of the surface from which the ions are extracted causes many of the trajectories to diverge from the desired paraxial direction, and to impact upon the accel or decel grids. The erosion that this causes severely limits lifetime[3].

This problem is avoided by the use of 4 grids, since the extraction field is defined by the first two of these

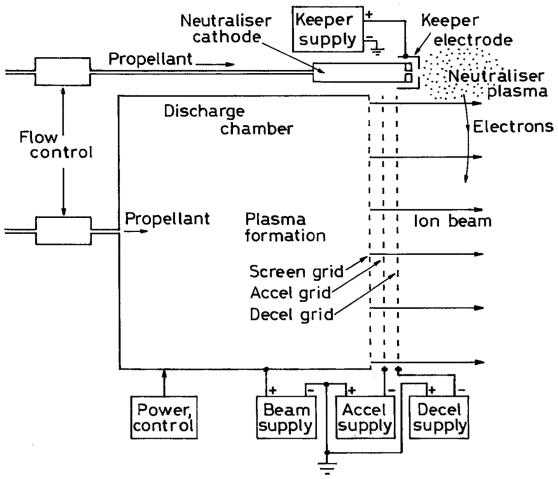


Figure 1: Schematic diagram of a gridded ion thruster

(Fig 3). Most of the ion acceleration then takes place between the second and third grids, where the greater part of the applied voltage appears. Thus the ion extraction process remains constant no matter what ion velocity and SI are required. This concept originated in the controlled thermonuclear research (CTR) community to produce very high energy particles for injection into fusion machines[7, 8, 9] and is used to accelerate ions, usually hydrogen, to energies that can exceed 100 keV.

Using standard ion engine scaling relationships[10, 11] coupled with the 4-grid extraction system, it can be shown that practical devices can have very high performances, which are suitable for interstellar precursor missions. For example, the performance of a 20 cm beam diameter thruster operating at relevant ion accelerating potentials is given in Table 1, assuming 50% of maximum grid perveance to ensure long life (this implies that one half of the theoretically maximum ion current density is extracted from the grids).

As an example of what has been achieved in the CTR community, the neutral injection machines developed and supplied by the UKAEA Culham Laboratory[7] have a rectangular grid area of 40 $\times$ 18 cm and operate with V_B of up to 80 kV. Using hydrogen, they achieve beam currents reaching 60 A, so that the beam power can be as large as 4.8 MW; this is a remarkable 6670 W/cm^2. The maximum thrust is 2.4 N (although it must be recalled that this is not operating as a thruster), and the SI is 400,000 s. As most of the power is deposited in the beam, the electrical efficiency is extremely high. The beam divergence is very small, at 0.4 °, suggesting that grid damage through primary ion impingement will be negligible. These and other data from CTR injectors are shown in Table 2; all employ a hydrogen plasma from which ions are extracted and accelerated.

3. Parametric Study of Mission Capabilities

It is clear from Table 3 that very high values of SI are necessary if the velocity increments required for interstellar precursor missions are to be achieved. Examples of the distances to be covered include 30 to 100 astronomical units (AU) to reach the Kuiper belt, 100 AU for the heliopause, and 550 AU for the gravitational lens focus of the sun. Of course, there is no possibility of covering the distance to the nearest stellar neighbours using this technology in timescales of interest. However, it was shown in this study that distances and mission times which can be termed interstellar precursor are viable using ion thruster technology and nuclear fission power sources.

The study assumed the use of the 4-grid ion extraction concept discussed above and the usual scaling relationships[10, 11] for the derivation of thruster operating parameters. While Xe propellant was adopted, there is no reason why another gas should not be used, provided that it does not react chemically with the thruster; in this way, a much wider range of values of SI is accessible. The rocket equation was employed to determine propellant mass, with ΔV = 35.7 km/s, to include escape from the Earth's gravitational field from a nuclear-safe initial orbit of 5000 km altitude[1] as well as escape from the solar system.

It is accepted that this relatively crude analytical approach will not yield accurate data, but it does allow an indication to be gained of what is feasible using this innovative technological approach. The above value of ΔV is, in fact, somewhat arbitrary; it was selected to be large enough for distances of interest to be covered in the 25 or so years mentioned previously. As will be discussed below, it is probably inadequate to escape fully from the solar system, since it ignores the gravity losses inherent in the proposed spacecraft trajectory. However, these losses are included in the ΔV assumed for the Earth escape phase of the

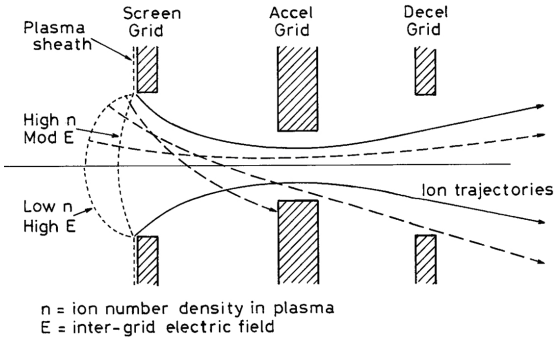


Figure 2: Triple-grid configuration, indicating the effects of the sheath shape on ion trajectories

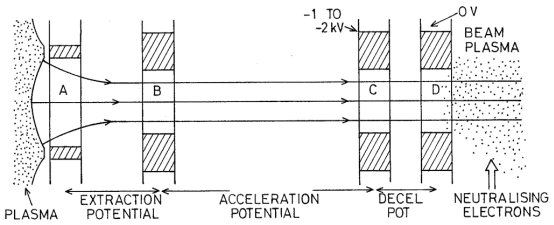


Figure 3: Schematic diagram of 4-grid system

Table 1: Predicted performance of a 20 cm beam diameter thruster utilising Xe propellant, a 4-grid ion extraction system operating at 50% perveance, and beam accelerating potentials of up to 70 kV

	Net Beam Accelerating Potential (kV)						
	7	10	15	20	30	50	70
Maximum thrust, 50% perveance (N)	2.4	3.3	4.1	4.7	5.8	7.5	8.8
Thrust density (mN/cm^2)	8.8	10.6	13.0	15.0	18.5	23.8	28.1
Beam current (A)	19.5	19.5	19.5	19.5	19.5	19.5	19.5
Input power at 50% Perveance (kW)	142	204	307	404	609	1013	1418
SI (s)	9300	11,100	13,600	15,700	19,300	24,900	29,400
Power density (W/cm^2)	451	644	966	1288	1933	3221	4509

Table 2: Characteristics of CTR machine ion beam accelerators (* Design Study ** Production devices from the UKAEA Culham Laboratory)

Authors	Grid Size (cm)	Grid Form	Open Area Ratio (%)	Beam Energy (keV)	Beam Current (A)	Beam Power (kW)	Beam Divergence (°)	Current Density (mA/cm^2)
Okumura[8]	10 dia	Flat	31	70	4-7	280-490	1.4	170-190
Ohara[12]*	12 dia	Flat	40	75	15	1125	0.6	133
Menon[9]	18 dia	Dished	51	35-65	7-20	245-1300	2	27-79
Martin[7]**	40 18	Flat	40	80	60	4800	0.4	200

Table 3: Typical velocity increments and values of SI for deep space missions

Mission	**Typical ΔV (km/s)**	**Minimum SI (s)**
Nearby planets	5-7	3000
Outer planets	10-15	3000-5000
100-1000 AU	100	10,000
10,000 AU	1000	100,000
Interstellar	30,000	3×10^6

mission.

An iterative procedure was required to assess tankage mass, assumed to be 6% of that of the total propellant load. Thruster masses were derived from existing devices, such as the T6[13], and the total propulsion system mass was designated M_{ps}. As it was assumed that the mission duration would be at least 25 years, and that a typical thruster lifetime is 20,000 hours if carbon grids are employed, it was decided that 12 thrusters would be required on the spacecraft, to provide both the total impulse needed and some redundancy. Gimbal thruster mounts were assumed to enable attitude control to be conducted by control of the thrust vectors.

In the initial phases of the study, the power source specific mass, α_n, was assumed to be 25 kg/kW; this is well within reach at the tens of kW level and higher, using turbo-machinery for power conversion[10]. The basic spacecraft mass, M_{sc}, was taken as 5 tonnes; this includes the instruments, computers and data storage, a very large antenna with a high power transmitter, thermal control, and so, together with the redundancy dictated by the long mission duration. The output of the power source was matched to the propulsion system, assuming a thrust, T, of 0.5 N, with most of the power being provided direct to the thrusters without intermediate converters. In this, a helpful characteristic of gridded thrusters is their insensitivity to relatively large voltage fluctuations[5].

Within these assumptions, the launch mass is plotted against SI in Fig 4, where the classical minimum in the curve is evident. This is at the optimum value of SI, which is about 20,000 s in this case. As shown in Fig 5, this corresponds to a value of V_B of approximately 35 kV and a total input power of 60 kW. From Table 1, these parameters can be provided by using a single 20 cm diameter thruster.

The study then proceeded to calculate, in a simplistic manner, the nominal distance which could be achieved within 25 years. In using the classic rocket equation[1] in the heliocentric phase of the mission, this calculation ignored the gravity losses which occur in a low thrust manoeuvre in a gravitational field. Basically, when the acceleration of the vehicle due to the onboard thrusters is considerably less than that caused by the local gravitational field, a much greater momentum transfer is required from the propulsion system than in the high thrust case. This can be represented as an effective loss of propellant mass. Thus further detailed work is needed to obtain more precise answers; however, the present very approximate approach does give an indication of what can be achieved by this proposed high SI, high power ion thruster technology. These approximate results are discussed below.

The distance achieved in 25 years according to this approximate method of calculation is plotted against SI in Fig 6 and the time to reach the specified ΔV in Fig 7. The former values are of interest in the interstellar precursor context, reaching beyond 140 AU, and the latter is achieved well within the 25 yr allowed. Since the basic spacecraft mass of 5 tonnes was somewhat arbitrary, the effect of varying this parameter was then assessed, adopting the range 2.5 to 10 tonnes, but keeping other factors at their previous values. The results are shown in Figs 8 and 9, from which it is clear that a vastly improved performance is available if this mass can be reduced. Significantly, the optimum SI then falls, the distance which can be achieved in 25 yr approaches 250 AU, and the time to escape from the solar system reduces to 10 yr.

The qualitative conclusions which arise from these results are not altered by the approximations mentioned above, although all times and distances are optimistic, owing to the gravity losses which are ignored in the calculations.

The effect of α_n on the SI optimisation procedure was then ascertained, with M_{sc} = 5 tonnes and T = 0.5 N. The results, shown in Fig 10 above, confirm the launch mass benefits of reducing α_n and indicate that the optimum SI increases as α_n falls, reaching about 24,000 s at 15 kg/kW.

Finally, the effect of thrust was examined, over the range 0.3 to 2 N. Assuming M_{sc} = 5 tonnes, α_n = 25 kg/kW and an SI of 20,000 s, which is near to the optimum value indicated in Fig 4, the results presented in Figs 11 and 12 were obtained. The launch mass and power consumption increase linearly with thrust, reaching about 13.5 tonnes and about 240 kW at 2 N. However, the time required to escape from the solar system reduces to about 7 yr and distances exceeding 300 AU are feasible.

4. Experimental Investigation at ESTEC

A proof-of-concept experiment was conducted in the CORONA vacuum facility of the ESA Propulsion Laboratory, located at ESA's research and technology centre, ESTEC, during November 2005. The principal aims were to demonstrate the practical feasibility of the 4-grid concept, to verify the high performance predicted by the analytical and simulation models, and to investigate critical design issues and technological challenges in preparation for any future spacecraft thruster development activities. This initial proof-of-concept experiment involved a small laboratory model prototype thruster, developed and built to ESA requirements by the Space Plasma and Plasma Processing Group at the Australian National University. During the test campaign, two different thruster configurations were successfully operated. The first was a single aperture four electrode (SAFE) configuration with a single 1 mm diameter aperture only, and the second was a dual-stage 4-grid (DS4G) variant incorporating 43 apertures of 1 mm diameter. Most of the beam tests were performed with the DS4G at higher thrust, once high voltage beams were shown to be extracted reliably from the SAFE design without direct ion impingement.

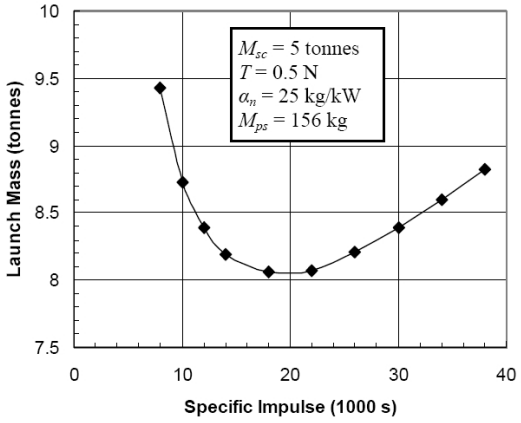


Figure 4: Launch mass as a function of SI for a payload of 5 tonnes

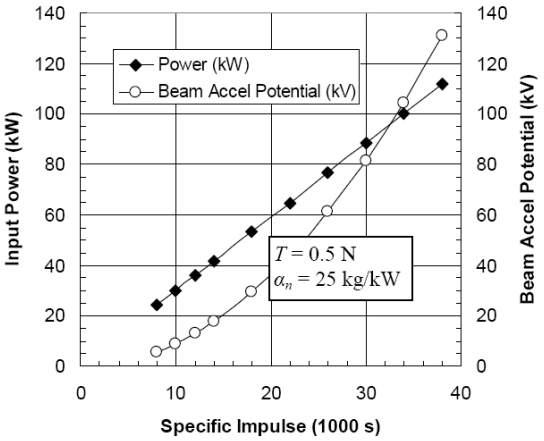


Figure 5: Input power and ion beam accelerating potential as functions of SI.

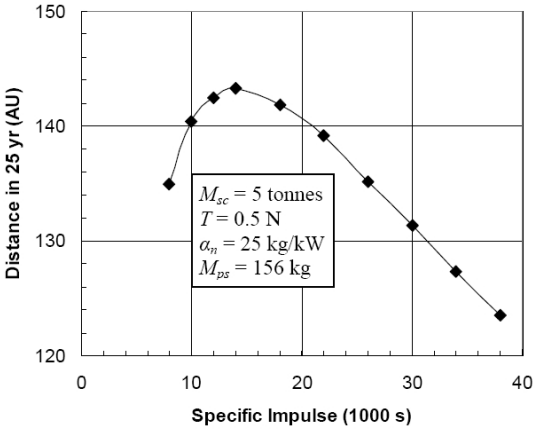


Figure 6: Nominal distance achieved in 25 yr as a function of SI

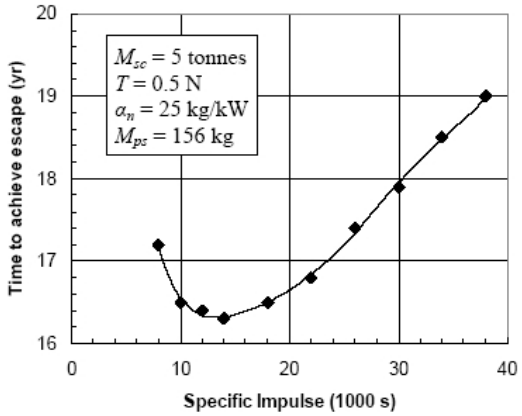


Figure 7: Time to attain velocity increment required for solar system escape as a function of SI

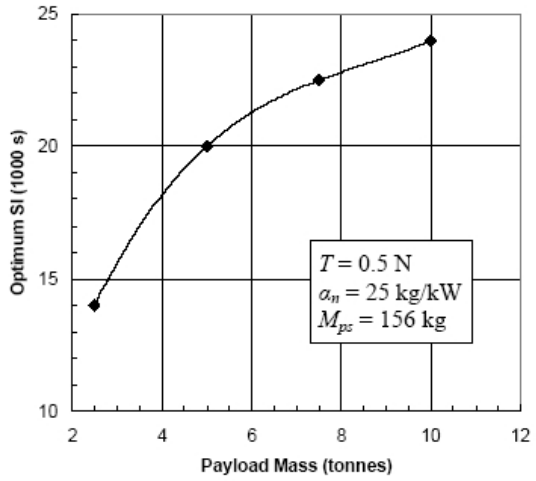


Figure 8: Optimum SI as a function of payload (basic spacecraft) mass

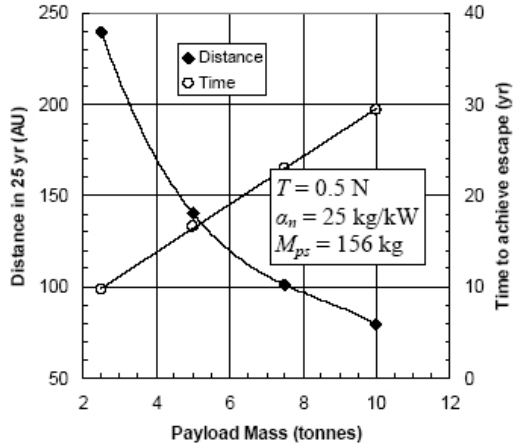


Figure 9: Nominal distance achieved in 25 yr and time to achieve escape velocity increment as functions of payload mass, at optimum SI

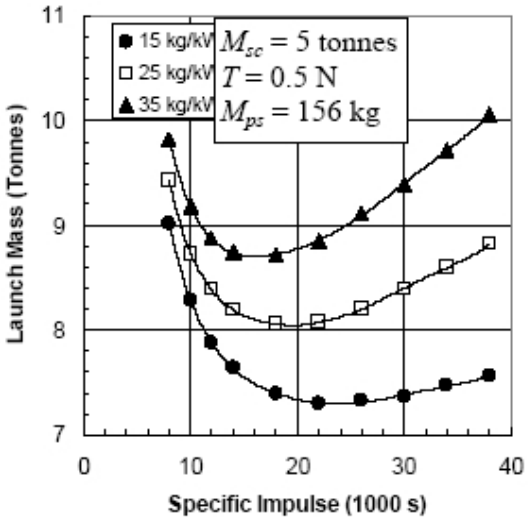


Figure 10: Launch mass as a function of SI for three values of power source specific mass

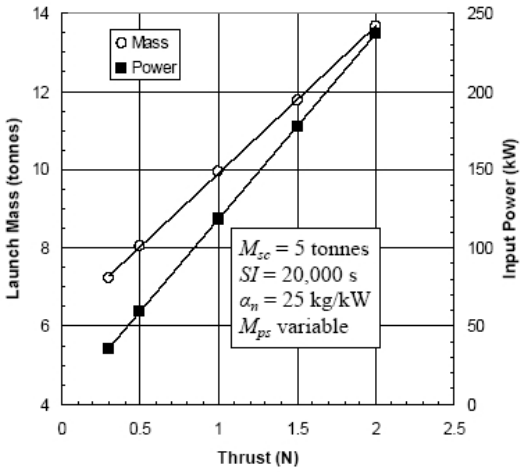


Figure 11: Launch mass and power input as functions of thrust

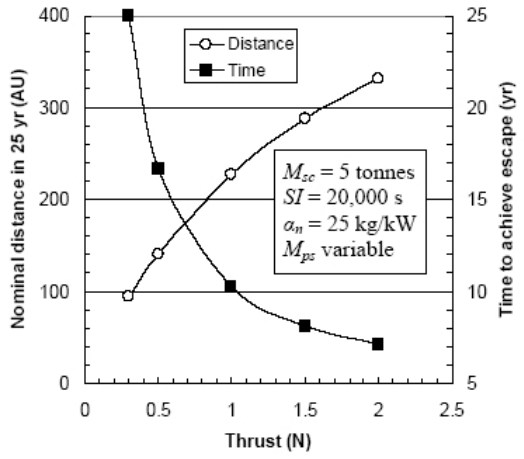


Figure 12: Nominal distance achieved in 25 yr and time to achieve escape velocity increment as functions of thrust, at optimum SI

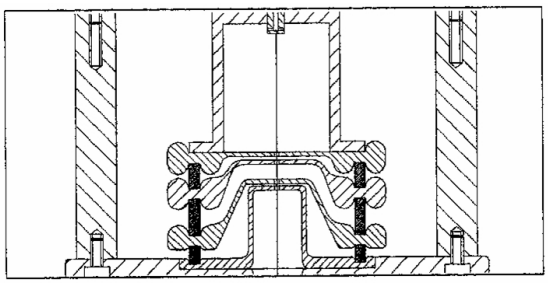


Figure 13: Sectional view of the thruster

Table 4: Summary of DS4G performance during the first tests in November 2005 at ESTEC

Maximum performance parameters		Operating parameters	
Thrust	2.85 mN	Beam potential	10 - 30 kV
Specific impulse	15,000 s	Extraction potential	3 - 6 kV (3 kV optimum)
Total efficiency	34 %	Beam current	4 - 12 mA
Mass utilisation efficiency	75 %	Beam power	100 - 260 W
Electrical efficiency	47 %	RF power	100 - 490 W
Beam divergence	2.5 - 5 °	RF discharge plasma density	2.5×10^{11} - 1.23×10^{12} cm^{-2}
Grid impingement	1 %	Mass flow rate	0.004 - 0.014 mg/s (Xe)
Thrust density (open area)	8.4 mN/cm^2	Beam diameter	2.3 cm
Power density (open area)	740 W/cm^2	Grid open area ratio	8.1 %
Thrust density (total)	0.7 mN/cm^2	Beam plasma electron temp	5.7 - 11.0 ($\pm$0.5) eV
Power density (total)	60 W/cm^2	Beam plasma potential	24 - 52 ($\pm$2) V

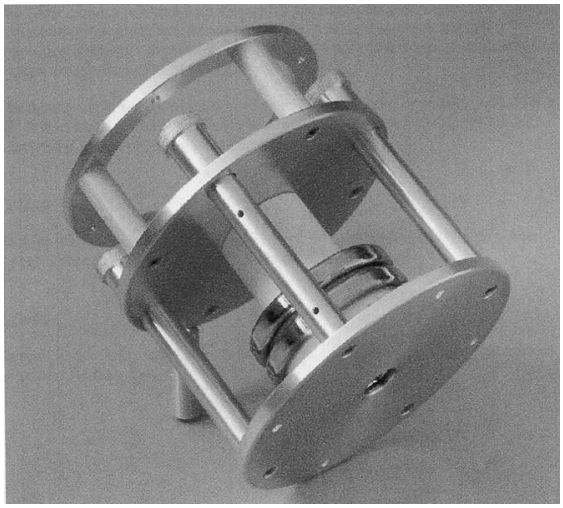


Figure 14: Photograph of the thruster

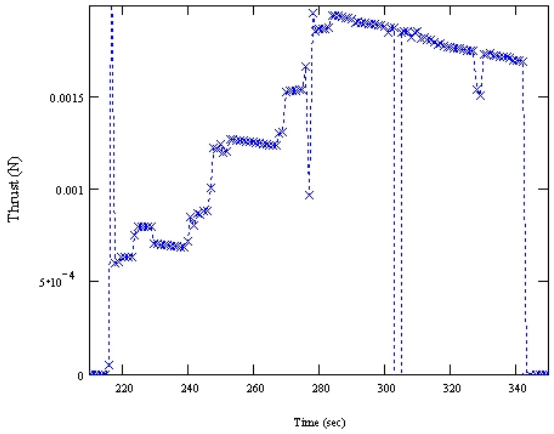


Figure 15: Thrust as a function of time for a 25 kV beam

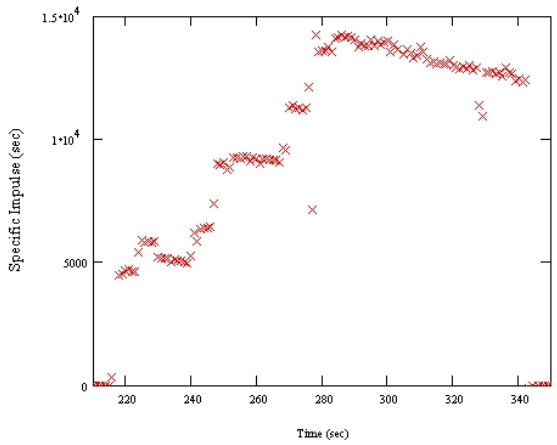


Figure 16: Specific impulse as a function of time for a 25 kV beam

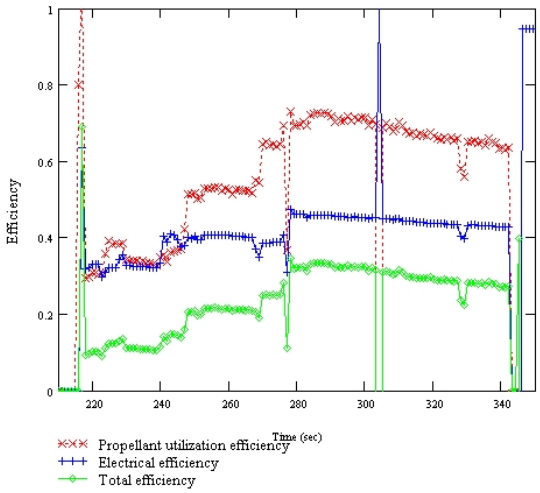


Figure 17: Thruster efficiencies as functions of time for a 25 kV beam

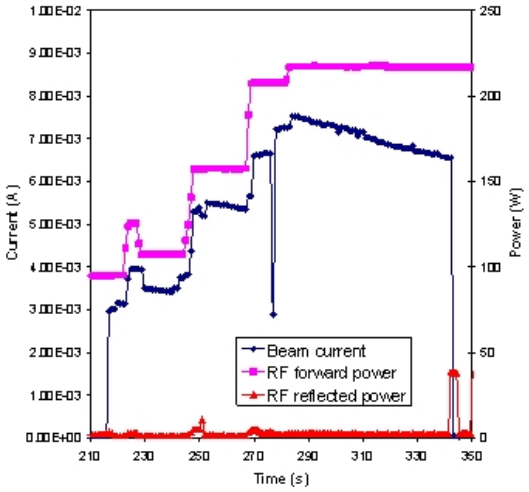


Figure 18: Beam current and RF power as functions of time for a 25 kV beam

The laboratory prototype thruster is shown in a sectional view in Fig 13, with the gas inlet at the top and the grid assembly at the bottom. A photograph is presented in Fig 14, in which the edges of the grids can be seen at bottom right. The thruster was based on a 5 cm diameter cylindrical discharge chamber made from alumina and fed with regulated flow of xenon gas. The radiofrequency (RF) discharge was energised by a 3-turn antenna coiled around the discharge chamber; this is not shown in Figs 13 and 14. This antenna was fed from a manually-tuned RF system operating at 13.56 MHz. A high voltage 4-grid system, connected to appropriate power supplies, was attached to the discharge chamber in order to extract ions from the plasma and accelerate them electrostatically via the two stage process discussed above. Under suitable conditions, this resulted in the desired highly focussed energetic ion beam.

The thruster was mounted and tested in the large 4 m long CORONA vacuum facility, with base pressures typically in the range 10^{-6} to 10^{-7} mbar. In total, some 66 different beam tests were performed on the SAFE and DS4G thruster configurations, with 54 tests for the latter version. Each beam test lasted for up to 10 minutes, which was long enough to acquire data on the grid currents and voltages, RF power and gas mass flow rate parameters. This information was sufficient to calculate the propulsive performance of the device. In addition, the characteristics of the extracted ion beam were investigated by using a suite of diagnostic tools. These diagnostics were operated at 1.5 m from the thruster and included scanning electrostatic wires for the determination of the beam divergence, and a Langmuir probe for obtaining the electron temperature and potential of the plasma surrounding the beam at various angles with respect to the thruster axis.

A summary of the test results from the DS4G thruster is presented in Table 4. The test programme was a remarkable success and the dual-stage 4-grid concept was proven to work in practice, in the context of space propulsion, for the first time. The thruster could be operated at beam potentials of up to 30 kV and produced an excellent performance in terms of specific impulse, thrust density and power density (in grid open area), and beam divergence. Despite the simplicity of the thruster and the very low open area ratio of the grids, the values achieved represent an improvement by several times on the current state-of-the-art, whilst maintaining negligible direct ion impingement of the beam on the grids.

The results of one typical beam test at 25 kV are shown in Figs 15 to 18. During this test, in which the ion extraction voltage was 3 kV, the RF power was increased in steps from 90 to 220 W, giving rise to immediate step increases in the beam current level from 3 to 7.5 mA. Beam current then decayed towards 6 mA, with the RF power maintained at 220 W. This caused a commensurate increase in thrust from 0.6 to 2 mN, decaying to 1.7 mN in steady-state. Occasional inter-grid arcing events occurred during the beam tests at these high voltages, causing the spikes in the plots shown. Since the mass flow rate was kept constant throughout this beam test, and was set to a level for a high expected beam current, the propellant utilisation efficiency was initially low at only 30%, but improved to 73% as beam current was increased by boosting the RF power. In turn, this improved the specific impulse to nearly 15,000 seconds at 7.5 mA.

Since the thruster was originally designed with one objective in mind, to prove the feasibility of the operation of the 4-grid concept at high beam potentials via the use of the SAFE configuration, the grid ion optics were not optimised and the open area ratio was very low, leading to a low beam power compared to the RF power. Therefore, the electrical efficiency and hence total efficiency achieved by the DS4G thruster were very low.

The need to improve efficiency to typical gridded ion

thruster levels before a flight engineering model can be contemplated has led to the initiation of a second test campaign utilising a revised grid optics design with a considerably increased open area ratio. This second test series will take place in May 2006 and simulation of the new grid ion optics indicates that beam currents of over 50 mA should be achieved (ie a factor of 10 increase) for the same RF power. This is expected to increase thrust to above 13 mN, beam power to greater than 1.6 kW and to raise total efficiency above 70%, for a slightly smaller beam diameter of 2 cm. It is also expected that specific impulse will increase towards the target of 19,000 s in the next test, through improvements in the propellant utilisation efficiency. Finally, this aspect and the addition of a high voltage isolator in the gas feed system should mitigate the arcing behaviour observed in the tests reported here.

5. Conclusions

It has been shown that an advanced form of gridded ion thruster, employing a 4-grid ion extraction and acceleration system, can provide thrusts, velocity increments and values of specific impulse of interest to interstellar precursor missions, with achievable distances extending to a few hundred astronomical units from the sun. In this it is assumed that a nuclear power source is available with a mass-to-power ratio in the range 15 to 35 kg/kW and an output of several tens of kW. Mission durations are of about 25 years and the velocity increment provided exceeds 37 km/s; the latter includes escape from the Earth's gravitational field after launch into a nuclear-safe altitude of 5000 km. It is shown that spacecraft basic mass, selected thrust level and power source specific mass have major impacts on mission performance. However, it should be noted that the numerical mission parameters derived in this study are optimistic, because the gravity losses inherent in low thrust heliocentric manoeuvres in gravitational fields have been ignored for simplicity.

The experimental programme conducted recently at ESTEC has demonstrated the viability of the 4-grid ion extraction and acceleration concept. For the RF plasma source employed, the optimum ion extraction potential was found to be about 3 kV utilising xenon propellant. Using this, total accelerating potentials of up to 30 kV were demonstrated, with a very narrow beam divergence of the order of 2°. The SI reached 15,000 s and the open area thrust and power densities 8.4 mN/cm^2 and 740 W/cm^2, respectively. These results are very encouraging, and indicate that more work to further develop the concept is fully justified.

Bibliography

[1] D.G. Fearn. The use of ion thrusters for orbit raising. *J. Brit. Interplan. Soc.*, 33:129–137, 1980.

[2] A plan to qualify a Russian space nuclear reactor for launch by the United States. In G. F. Polansky, G. L. Schmidt, E. L. Reynolds, E. D. Schaefer, B. Ogloblin, and A. Bocharov, editors, *AIAA, SAE, ASME, and ASEE, Joint Propulsion Conference and Exhibit, 29th, Monterey, CA, June 28-30, 1993, 12 p.*, June 1993.

[3] A.R. Martin, C.R. Banks, R.W. Baton, P. Hurford, W.B.F.L. MouIford, and A.J. Pearce. Erosion measurements for two- and three-grid ion thruster extraction systems. IEPC Paper 93-171, September 1993.

[4] D.G. Fearn and S.W. Patterson. Characterisation of the high current hollow cathode for the t6 ion thruster. In *AIAA, SAE, ASME, and ASEE, Joint Propulsion Conference and Exhibit, 34th, Cleveland, OH*, number AIAA Paper 98-3346, July 13th-15th 1998.

[5] D.H. Mundy and D.G. Fearn. Throttling the t5 ion engine over a wide thrust range. In *AIAA, SAE, ASME, and ASEE, Joint Propulsion Conference and Exhibit, 33rd, Seattle, WA*, number AIAA Paper 97-3196, July 6th-9th 199.

[6] R.A. Bond, D.G. Fearn, N.C. Wallace, and D.H. Mundy. The optimisation of the uk-10 ion thruster extraction grid system. IEPC Paper 97-138, 1997.

[7] A.R. Martin. High power beams for neutral injection heating. *Vacuum*, 34(1-2):17–24, 1984.

[8] Y. Okumura et. al. Quasi-DC extraction of 70 kev, 5 a ion beam. *Rev. Sci. Instrum.*, 51:728–734, 1980.

[9] M.M. Menon et. al. Power transmission characteristics of a two-stage multiaperture neutral beam source. *Rev. Sci. Instrum.*, 51:1163–1167, 1980.

[10] D.G. Fearn. The application of gridded ion thrusters to high thrust, high specific impulse nuclear-electric missions. IAF Paper IAC-04-R.4/S.7.09, October 2004.

[11] D.G. Fearn. The design of a versatile ion thruster for large communications spacecraft. IAF Paper IAF-97-S.3.01, October 1997.

[12] Y. Ohara. Numerical simulation for design of a two-stage acceleration system in a megawatt power ion source. *J. Appl. Phys.*, 49:4711–4717, 1978.

[13] J. Huddlestone et. al. An overview of the t6 gridded ion propulsion system pre-development activities for alpha-bus. In *Proc 4th International Spacecraft Propulsion Conference*, Sardinia, 2-4 June 2004.

Formation Flying Picosat Swarms for Forming Extremely Large Apertures

Ivan Bekey

President, Bekey Designs, Inc. 4624 Quarter Charge Dr. Annandale, VA 22003, USA; ibekey@cox.net

Abstract

This paper serves as a brief introduction to forming a very large array antenna in space. Some reasons why such large antennas are desirable are stated, followed by reasons why beyond some size they must consist of sparse arrays of very small spacecraft, whose signals are coherently combined to form a very large yet lightweight antenna, the requirements for coherent addition are discussed, and some aspects of the antenna pattern are presented. Some orbit design considerations are discussed, and the questions of where and how to locate the feed are discussed with reference to sensitivity to spacecraft position knowledge. Some considerations of how to maintain formation are also addressed.

1. Introduction

Very large aperture antennas could find a number of intriguing applications in a variety of orbit altitudes, as the mainbeam angular resolution is directly proportional to antenna diameter, and its sensitivity is proportional to the square of the diameter. Such applications could include earth resources sensing, communications, radar, and others, in which the ground spot diameter needs to be as small as possible for either sensitivity or directivity reasons. Nonetheless the largest antenna in use today is the Thuraya communications satellite at 12.25 meters diameter, and launches are planned of a 22 meter antenna communications satellite, the sizes limited principally by the weight and cost of the antennas. In communications applications the ground footprint of such antennas will be limited as shown in Figure 1.

It is seen that in order to attain mainbeam footprints of, for example, 10 km at a frequency of 1 GHz from GEO an antenna diameter of 1,200 m would be needed. If it were placed into LEO its size would need to be "only" 120 meters (not looking straight down but rather to the horizon).

Admittedly most communications applications aim for large earth coverage, but some emerging applications do not. In applications requiring high sensitivity as well as large coverage they are forced into generating very many simultaneous beams in order to attain large coverage and simultaneously the high gain resulting from small spot size. In the latter case the sizes above hold. These are very large antennas indeed, and quite beyond the state of the art.

2. Very Large Antennas

Antennas of hundreds, let alone thousands of meters in diameter, cannot be implemented using any technique known today if they are to be filled apertures, as Figure 2 illustrates.

Clearly the weight of a filled aperture of that size is astronomical, no matter what its construction technique, even considering fully inflatable membrane antennas which are not yet state of the art, and even if they were they would be too heavy and costly by many orders of magnitude.

From the foregoing it is clear that only a very sparse array with no structure *per se* can be used to implement a very large antenna. In principle, all that is needed is to disperse antenna elements widely separated in a constellation in space. But this means that the usual way of feeding the elements by waveguides to each element cannot be implemented, and the array must be remotely fed, or what is termed "space-fed". This will be expanded upon later, and is illustrated in Figure 3.

The antenna must be implemented as many small elements dispersed in a plane in space, with the area occupied by

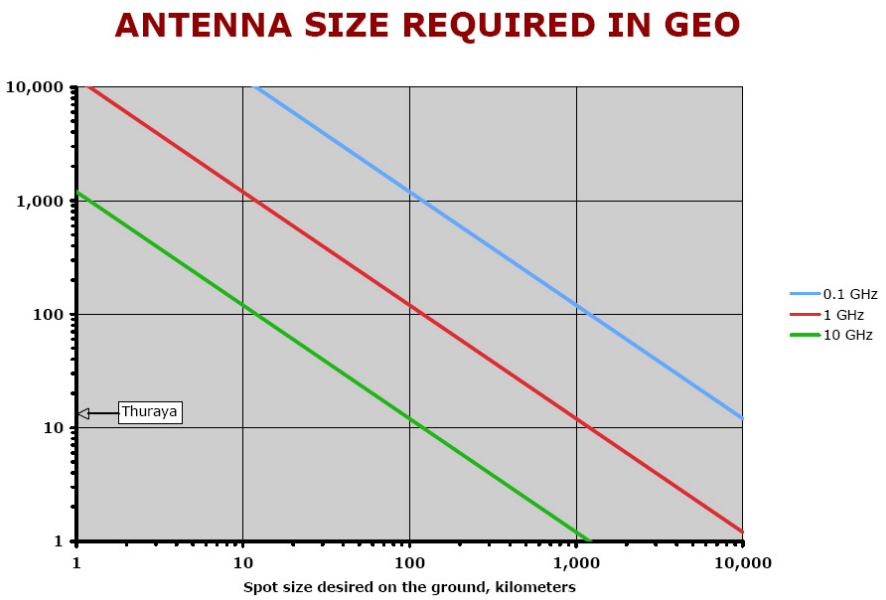


Figure 1: Antenna size in GEO required to form a spot of a given size on the ground

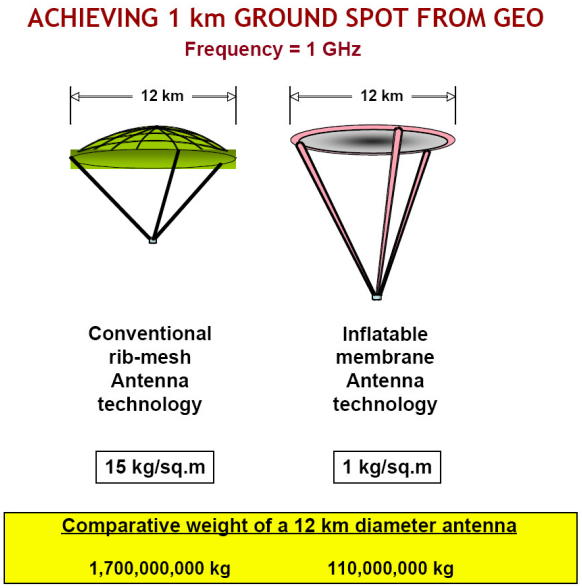


Figure 2: Weight of a 12 km diameter antenna

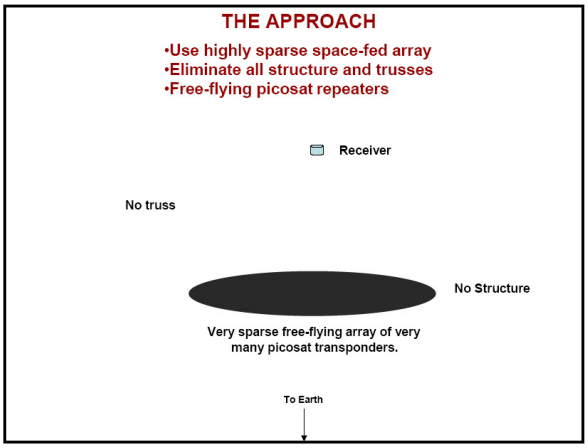


Figure 3: Space fed structure-less sparse array

the elements being a very small portion of the total area of the constellation. The feed is located above the array, as shown, or it could be on the ground. The size of the array is thus independent of its collecting area, and the lack of any structure makes for a very lightweight configuration. As an example, an array consisting of 1,000 elements, each of which is configured as a self-contained picosat weighing 1 kg, would have a total weight of 1,000 kg regardless whether the constellation diameter were 100 m or 100 km.

3. Orbit Considerations

Dispersal of many elements of an antenna in space into a free-flying constellation requires either that the relative positions of the elements be immaterial to antenna performance, or that those positions be controlled by some sort of formation flying so that the antenna performance can be invariant with time.

We will assume for the remainder of the discussion that the antenna is a receiver whose elements are contained in individual picosats which act as phase-controlled repeaters to receive the energy from the ground and retransmit it toward a central receiver to be added coherently. This is the space-fed sparse array whose design parameters will be discussed in the remainder of this short paper. Furthermore we will assume that there need to exist many picosats, in the order of 1,000 or more in the array, and that the array needs to be 1-100 km in diameter to attain a desired small ground spot size at a low microwave frequency. These numbers are not unreasonable given some applications of communications and remote sensing which are very desirable to conduct from GEO due to the constant coverage of particular areas.

The picosats need to be in a constellation that is both controlled and requires little propellant to maintain the relative positions of the picosats. While it is possible to hold picosats in well separated, relative positionsin earth-centered coordinates, say along the local vertical and local horizontal, and thereby form a static constellation antenna array, the propellant requirements for a kilometers-across antenna array for such station keeping in non-Keplerian orbits rapidly become excessive. As an example, the mass of propellants required for propulsively stationkeeping a 1 kg picosat 50 km above or 16.7 km cross track from the Keplerian GEO orbit for 10 years is 8.25 kg, even using a propulsion system with Isp of 3,000. This is more than 800% greater than the basic picosat mass assumed, and is clearly prohibitive.

A better choice is to place the picosats into a set of orbits obeying Hill's equations in which they describe Keplerian orbits, but that rotate in apparent sub-orbits around a central point on the base orbit when viewed in relative coordinates centered on that point. Since such orbits are Keplerian, the picosats' propulsion requirements are minimized.

These so-called HALO orbits are illustrated in Figure 4.

In these equations a particular family of solutions, developed by C.C. Chao of The Aerospace Corporation, are discussed. A family of solutions with no constant and secular (in time) terms can be found:

$$
\begin{aligned}
x &= 2\xi z_0 \sin \omega t, \\
y &= -\xi z_0 \cos \omega t, \\
z &= z_0 \cos \omega t, \\
\frac{dx}{dt} &= -2\omega\xi z_0 \cos \omega t, \\
\frac{dy}{dt} &= -\omega\xi z_0 \sin \omega t, \\
\frac{dz}{dt} &= -\omega z_0 \sin \omega t
\end{aligned}
$$

The parameter ξ is a scale factor which gives a family of solutions for sub-orbits, and the magnitude of the radius vector, r, of the sub-orbit around the center satellite is

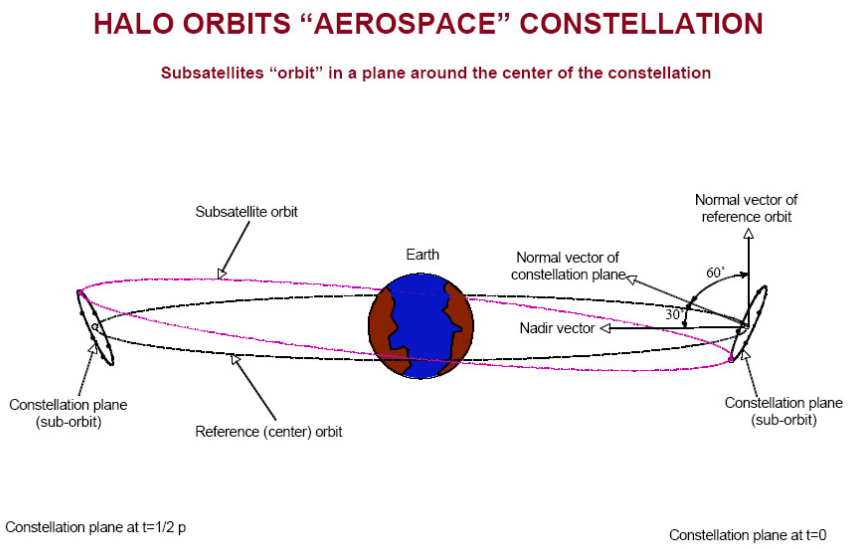


Figure 4: HALO orbits obeying Hill's equations

$$r = (x^2 + y^2 + z^2)^{1/2}$$
$$r = z_0[4\xi^2 + (1 - 3\xi^2)cos^2\omega t]^{1/2}$$

Among the infinite number of solutions that depend on the value of ξ, only two particular values are of interest in this application. The above equation indicates that the sub-orbit is a circle when ξ equals (1/3)1/2. The second value of ξ is $\frac{1}{2}$, which would make the sub-orbit an ellipse whose projection on the plane normal to orbit plane is a circle.

Through mathematical derivations (vector products), one can prove that the sub-orbits are in a plane and the angles between the normal vector of the plane and the z axis (Earth spin axis) are $\delta = arctan(1/\xi)$, which is 60 deg when ξ = (1/3)1/2 and 63.4 deg when $\xi = \frac{1}{2}$. These angles, of course, mean that the sub-orbit plane will lie at 30 or 26.6 degrees to the local horizontal at the base orbit. Of the two values of ξ, the choice of 26.6 degrees is preferred because it gives a halo sub-orbit with a constant radius in its orbit around the central point. Furthermore a number of previous studies have shown that a circular halo sub-orbit in its plane requires less propellant for constellation maintenance.

The orbital parameters of different diameter HALO constellations in the GEO orbit are shown in Table 1.

The appearance of such a space-fed sparse array antenna is shown in Figure 5.

The many picosats (100-10,000) may be populated in multiple concentric rings or in multiple spiral arms or other distributions in the sub-orbit plane. Based on one first-cut example system design, an array containing 12,000 picosats can be uniformly populated on concentric rings with largest diameter (the constellation diameter) being 100 km. This constellation requires 31 concentric rings at 1 km spacing in radius.

The spacing along the circumference direction is also set at 1 km. The innermost ring has a diameter of 20 km which allows more than enough room for the central tether that could be used to hold the central receiver without need for constant propulsion to maintain position relative to the array.

The three types of natural forces dominating the orbit geometry at geostationary altitude are the Earth gravity harmonics, luni-solar attractions, and the solar radiation pressure. These perturbing forces gradually push the picosats away from their design orbits, both individually and as an ensemble. Numerical results from past analyses show that the ensemble position deviation in near-GEO would exceed 100 m in less than two hours, 1000 m in little over 5 hours and 10 km in about 12 hours. However, the relative distance between two neighboring picosats remains nearly the same even after 10 days, due to the nearly equal perturbing forces on the two closely separated picosats.

Nonetheless, analyses and video simulations reveal that, without orbit control, the relative positions of the picosats begin to change rapidly after 10 days causing separation changes that would be undesirable for array pattern reasons, as well as causing collision dangers and perhaps even loss of picosats which would become orbital debris. Thus even in a HALO constellation the picosats will require propulsive translation control for formation flying, however the total impulse requirements per picosat will be an order of magnitude lower than if the constellation geometry were maintained in the same constellation wholly by engine thrust in non-Keplerian orbits.

In fact estimates indicate that the velocity change requirements on any one picosat are substantially less than

Table 1: Orbit parameters for different size constellations in GEO

Halo ring Diameter	Semi-major axis	Eccentricity	Inclination	Ascending node	Argument of perigee	Mean anomaly
100 km	42165.9 km	0.0005926	0.05884 deg	0°	270	20
98 km	42165.9 km	0.0005816	0.05766 deg	0.5844°	270	199.415
20 km	42165.9 km	0.0001186	0.01177 deg	21.675°	270	178.32

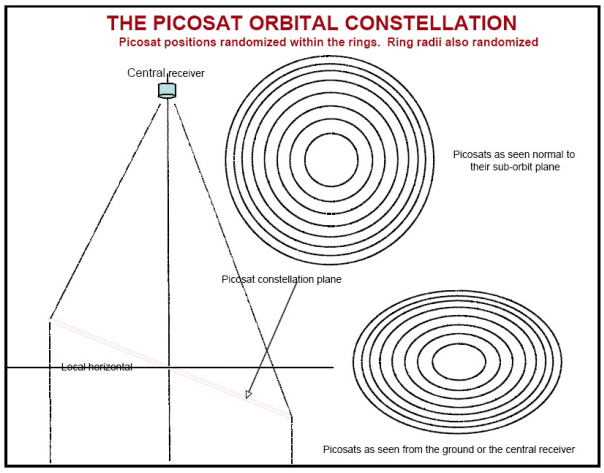


Figure 5: The HALO orbit constellation

100 m/sec per year and may actually be on the order of 5-10 m/s per year. While this is clearly more attainable, particularly in highly integrated and compact picosats, the tracking and control of thousands of separate spacecraft in relative proximity is far beyond anything tackled today, since principally 2-body problems have been addressed to date, not 1,000-body problems.

In practice the picosats are deployed each into a slightly different orbit, and fly in circles around the GEO point in relative coordinates. They appear to be glued to a transparent plane tilted at some angle to the local horizontal, the preferred angle being 26.6 degrees, for that results in circular sub-orbits and also minimizes the propellant expenditure to stationkeep the picosats.

This angle is patented by The Aerospace Corporation, USA.

4. Functioning of a Space-Fed Array Antenna

The large constellation array envisioned is thus comprised of a set of picosats in a plane, whose prime requirement is for each picosat to receive the signal from the ground, shift its frequency, and relay it to a central receiver where all picosat signals are added coherently, while their noise is not. The combined signals thus will be N times the signal of that of any one picosat, while the noise only adds non-coherently, or as the square root of the number of picosats. This means that there is a square root of N advantage to having a large number N of picosats in the array.

In a space fed array the picosats act as refractive elements akin to a lens in optical systems. The advantage of a space-fed array with central receiver above the array plane is that, in principle, position errors of the picosats do not, to a first order, cause phase errors in signals repeated by the picosats because the sum of the path lengths from the ground to the picosat and from the picosat to the central receiver remain constant with picosat position changes.

The space-fed sparse array functioning is illustrated in Figure 6.

Each picosat receives a small portion of the essentially plane wave arriving from the ground source. The job of the picosats is to repeat those signals toward the central receiver, each picosat changing its phase in the process. The purpose of that change is to convert the incoming plane wave into a spherical wave converging on the central receiver where all picosat signals will be added coherently. In order for the ensemble of picosats to achieve that result, the wave passing through the center of the array must be delayed relative to the wave at the edge of the array, since the latter wave has further to travel. If the system bandwidth were large, true time delay would be needed in the picosats to prevent dispersion. However for the relatively narrow band operation assumed here, phase correction suffices. A happy side benefit is that since two sine waves cannot be distinguished between phases 2π apart, the phase corrections can be done by advancing or by delaying the phases, and a 2π correction range

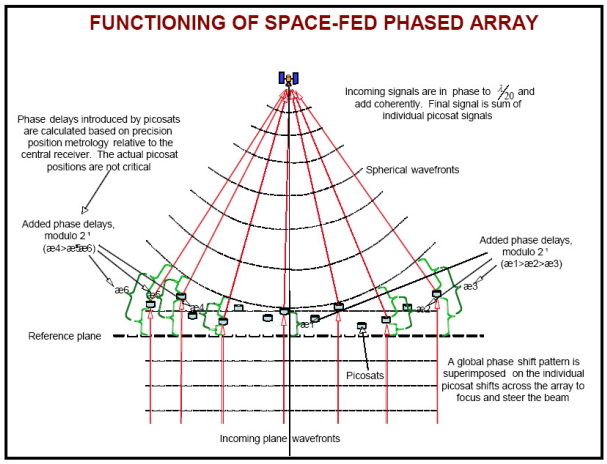


Figure 6: Functioning of a space-fed array

covers all possible phase corrections. This is referred to as Modulo-2π phase delay.

Now consider the actual corrections that will have to be applied by each picosat. First note that for coherent addition we generally would like the phases of all signals to arrive within about 1/10 of a wavelength from each other. At a frequency of 3 GHz the wavelength is 10 cm and so the tolerance will be about 1 cm in position. Thus the position of each picosat must be known to about 1 cm for phase control, but its actual position control can be much looser, and will be set by constellation orbit pattern maintenance and other considerations, not by phase control reasons.

Thus a space-fed array acts as a corrective lens in space and does not require that the picosats comprising it be in precise locations 1/10 wavelength from a perfect plane in space. In addition to the corrections accounting for the actual picosat position deviations from a perfect plane, a global phase shift pattern can be imposed on the picosats, over and above the position control-dictated phase delays, in order to steer the beam and focus the antenna. These global phase patterns will probably be commanded from a central control source collocated at the central receiver.

A very important additional function must be met by the phase control - the compensation for the 26.5 degree tilt of the constellation plane with respect to local horizontal, which must be compensated by a global phase shift across the array. In effect the beam must be squinted 26.5 degrees constantly to point to the nadir, upon which is superimposed the steering phase shift program for dynamic beam steering as well as the individual picosat position-related phase shifts.

Now let's derive the actual phase control needed. Consider the picosat on the axis of the array. The total path length to the receiver consists of the path from the ground to the picosat plus that from the picosat to the central receiver. If the picosat should move upward along the axis, the ground-picosat path will lengthen while the picosat-receiver path will shorten. However, on axis the sum of the path lengths is constant regardless of the picosat position. Thus the space fed array is insensitive to picosat position on axis, and in principle would not require any phase correction. This is illustrated in Figure 7.

This complete insensitivity is limited to on-axis picosats, of course. However a similar calculation for an axial position error of a picosat at the edge of a 100 km array results in a sensitivity of 45 m motion before the phase error accumulates to the 1/10 wavelength. Similarly an analysis has shown that position errors in the sub-orbit plane are no more sensitive in terms of causing path length changes for on-axis picosats, however for motions in plane at the edge of the array the picosats can move only 28 cm before the phase errors add to 1/10 of a wavelength.

This demonstrates one of the benefits of using a space-fed phased array with receiver above the array plane compared to having the central receiver on the ground, in which all picosat positions must be determined to 1/20 of a wavelength, or $\frac{1}{2}$ cm, regardless of where or in which direction they move because the total path length will be double the excursion rather than less–that is its sensitivity is at least twice that of the array with tethered receiver.

In the above we have shown that that the position of the picosats must be known to 1/10 of a wavelength, or about 1 cm at 3 GHz as a worst case. The positions of the picosats can be determined either remotely, say by a radar or lidar in the central receiver, and transmitted to each picosat so its transmissions can be properly phase shifted; or the picosats can each determine its own position using GPS, or by setting up a dedicated local navigation environment consisting of a number of reference spacecraft in HALO orbits around the constellation array. Either technique is inherently able to provide the necessary accuracy and will enable the necessary phase corrections.

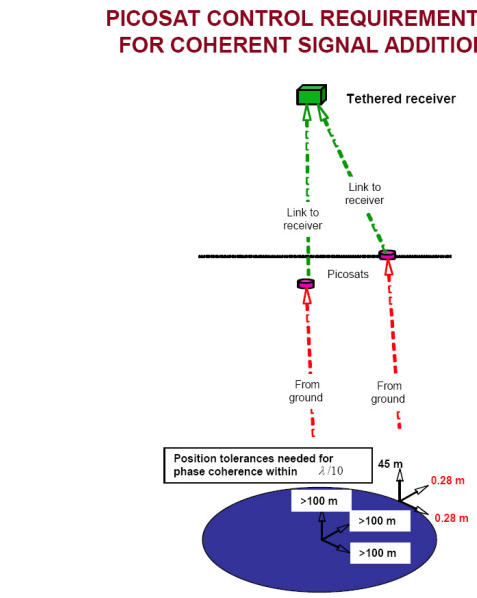


Figure 7: Picosat position knowledge requirements

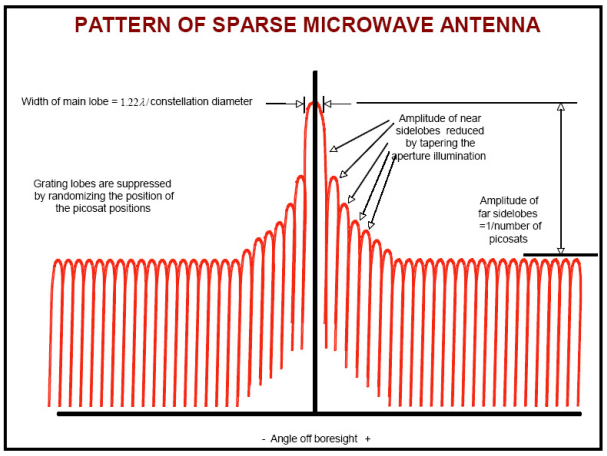


Figure 8: Sparse antenna pattern

5. Array Antenna Pattern

Sparse arrays behave differently than conventional filled arrays, in some very important ways. The pattern of a typical sparse array is illustrated in Figure 8.

The mainbeam width of a sparse antenna is identical to that of a filled aperture antenna of the same overall diameter. The grating lobes that are associated with use of a finite number of elements can be essentially eliminated by randomizing the relative positions of the picosats in the array. The sidelobe widths and number of sidelobes are also essentially the same. The amplitude of the near-in sidelobes are also similar, and governed by the amplitude taper (illumination) of the overall array, which can be controlled by adjusting the amplitudes of the transmissions from each picosat.

The main difference is that whereas in a filled aperture the far sidelobe amplituders are monotonically decreasing in amplitude as their angular distance from the mainlobe increase, in a sparse array their average amplitude is fixed and equal to 1/(number of picosats) regardless of their angular distance from the mainbeam.

This is a very important difference, as interference through the sidelobes, which might be tolerated or ignored for a filled aperture, may be dominant in sparse apertures due to their very much greater average sidelobe amplitude. The only remedy is to increase the number of elements (picosats) in order to reduce the sidelobe amplitudes relative to the mainbeam, increase the antenna aperture of each picosat in order to superimpose its filled pattern on the sparse pattern, or both.

In addition, since the total receiving area of the sparse array is the sum of the capture areas of the picosats, to obtain a large capture area there would have to be very many picosats, or the receiving antenna of each picosat must have an aperture larger than that of a simple dipole to keep the number of picosats reasonable.

Thus the sparse antenna, though having to have a large number of picosats, each of which may require some filled aperture receiving antenna, has the major benefit that its mainbeam width and total aperture area can be independently controlled, something not possible with filled apertures.

6. Conclusions

This brief paper has discussed sparse space-fed antenna arrays consisting of very many independent picosat elements for forming extremely large antennas in space.

Formation flying will be a necessity in all such large arrays, and placing the picosats into HALO orbits will minimize picosat translation propulsion requirements. The peculiar characteristics of sparse arrays must be understood in order to apply them correctly, nonetheless it seems evident that antennas tens, if not hundreds, of kilometers in size can be implemented this way, something not possible with any filled aperture antenna technologies known or even projected.

Agile Maneuvers for Near Earth Object (NEO) Fly-by Missions

Vaios Lappas[1], Bong Wie[2] and Jozef van der Ha[3]

[1] Surrey Space Centre, University of Surrey, GU2 7XH, United Kingdom, E-mail: v.lappas@surrey.ac.uk
[2] Arizona State University, Tempe, AZ 85287-6106, USA, E-mail: bong.wie@asusat.edu
[3] Consultant, 10001 Windstream Drive, Columbia, MD, USA, E-mail: JvdHa@aol.com

Abstract

Small Satellites are establishing themselves as important tools for exploring our solar system. Developments in micro-electronics have enabled small and low-cost deep space probes to complement conventional space platforms in long-duration deep space missions. The paper presents a baseline design of an attitude control system for a spacecraft performing a Near Earth Object (NEO) fly-by mission. Following the autonomous escape and cruise phases, the NEO encounter phase is the most critical mission phase with the attitude control performed by a cluster of small Control Moment Gyros (CMG's). The satellite needs to be rotated relatively fast to keep the NEO within its field of view for the imaging of the NEO. Simulations demonstrate the practicality and versatility on the use of CMGs for low-cost NEO deep space missions.

Introduction

It is known that Near Earth Objects (NEO's) pose a potentially catastrophic danger for earth. Thus there is a need to better understand NEO's and to better predict their orbits. Low-cost deep space probes can be useful and cost-effective for gathering information on NEO's. A mini-satellite mission that is capable of supporting a 10-kg science payload will be presented here. The main mission objective is to demonstrate the capability to intercept a NEO in deep space and to perform surface imaging. Surrey Satellite Technology Ltd. (SSTL) has a proven track record of successful low-cost satellite missions and has ambitions to demonstrate the same cost-effective design philosophies for deep space applications [1, 2]. The conceptual platform design assumed in the present study is basically identical to the one resulting from a previous SSTL system design study for an 'entry level' NEO mission [3]. The present paper focuses on the design of the Attitude Determination and Control System (ADCS) for a low-cost SSTL-built mini-satellite performing a NEO flyby. The proposed concepts are illustrated using the example of a low-cost mini-satellite that performs an imaging mission to NEO 4179 Totatis shown in Fig. 1. The low-cost nature of the mission dictates that, during the fast fly-by phase, the satellite must execute a fast rotation about its pitch axis in order to be able to keep its imager pointing at the target object.

The proposed ADCS design uses small agile Control Moment Gyro (CMG) actuators and low-cost attitude sensors. The Micro-CMG's provide the fast spacecraft rotation rate (up to 6 deg/sec) that is needed to perform the NEO imaging during the critical encounter phase.

2. Mission Analysis

2.1 Selected NEO Target

On the basis of the results of previous studies [3, 5] performed at the Surrey Space Centre, a single suitable candidate NEO was selected for the present study: 4179 Toutatis (Fig. 1) which is relatively large and a potential 'contact binary'. It came within a scant 1.5 million kilometers of the earth on 29 September 2004 and will approach earth again in 2008. The Toutatis object is of interest because it achieves one of the closest earth approaches of any known asteroid or comet between now and 2060, and the approach occurs in the near term.

Figure 1: Radar Image of 4179 Toutatis [4])

2.2 Trajectory

There are a number of potential low-cost launch opportunities, but the lowest cost launch is most likely offered by the PROTON rocket, which provides a secondary payload capability of about 400 kg into a geo-stationary orbit. When starting from this relatively attractive orbit, earth escape can be achieved at a cost of about 1300 m/sec. The encounter with 4179 Toutatis in 1998 would require an additional 200 m/sec [3]. The proposed NEO2M (NEO Mini-satellite Mission) trajectory consists of three parts: (i) the 267-days escape phase, when the satellite escapes from the earth's sphere of influence starting from its initial geostationary orbit; (ii) the 134-days approach phase when the spacecraft cruises to the NEO Toutatis; (iii) the short encounter phase when it conducts the actual NEO flyby. During the escape phase the semi-major axis is increased by means of relatively short perigee kicks. At the time of earth approach, Toutatis is near its perihelion and has a speed of about 40 km/s. The relative speed of the mini-satellite and Toutatis at encounter is close to 10 km/sec.

2.3 Imaging

Currently, existing Toutatis imagery is limited to radar data at relatively low resolution (Fig. 1), collected at JPL's Goldstone site. The collection of optical imagery would indisputably be a crucial objective of the Toutatis mission. The availability of 10-meter or better optical images would greatly increase our knowledge of this particular NEO and would improve our understanding of its complicated dynamics and structure. The NEO2M satellite will have two imagers: the Wide Field Of View (WFOV) and the Narrow Field Of View (NFOV) cameras. The WFOV imager has a 20° by 20° field of view and allows target acquisition from about 250,000 km away at 7 hours before closest approach. The WFOV camera produces wide-angle NEO images and provides inputs for optical navigation. The proposed NFOV camera is a miniaturized micro-imager derived from the SMART-1 camera. It has a 4° x 4° field of view, a 1024 x 1024 detector array, and a 120 mm main lens for capturing large NEO surface features. The NFOV imager is capable of acquiring about 40 seconds worth of high-resolution images at a rate of 30 frames per second. A more manageable concept (from a low-cost design point of view) may be based on five NFOV frames per second for a total of 200 images. The NFOV camera should be capable of achieving a GSD resolution of less than 10 meter around the time of closest approach with Toutatis (with length of 6.5 km) almost completely filling the imager's field of view. Although it may well be feasible to enhance the present capabilities of the NFOV micro-imager by using higher-resolution detector arrays, only the existing design is considered in the present paper. The 4° field of view leads to a GSD resolution of 6.8 m per pixel at the minimum miss distance of 100 km. On the basis of this design baseline, we find that NEO Toutatis fills less than 2% of the total image at the time when the probe is 5000 km away from the object. When we assume a worst-case system-level pointing capability of about 0.1 deg, we see that this corresponds roughly to the size of Toutatis at a 5000 km distance. At closest approach, the assumed pointing error amounts to about 175 m or less than 3 $\%$ of the NEO size. It can be expected that, at the subsystem level, the largest part of the total pointing error budget can be allocated to the ADCS design. In terms of pointing stability requirements, simulations indicate that 3×10^{-3} deg/sec would be sufficient for the present mission objectives.

2.4 Required Tracking Rate

Fig. 2 shows the (planar) encounter geometry of the probe relative to the NEO near the time t_{CA} of closest approach. D denotes the miss distance at the time t_{CA} and $R(t) = V(t - t_{CA})$ represents the varying distance between the probe and the NEO as a function of time. V is the relative flyby velocity which may be assumed to remain constant over the short interval of time under consideration. The required probe's pitch rotation angle equals the NEO aspect angle $\alpha(t)$ as shown in Fig. 2. In order to ensure that the camera keeps pointing towards the NEO during the encounter phase, the pitch angle profile that must be tracked by the ADCS is:

$$\alpha(t) = \arctan\left\{\frac{D}{R(t)}\right\} \tag{1}$$

Fig. 3 shows the evolution of $\alpha(t)$ near the time of closest

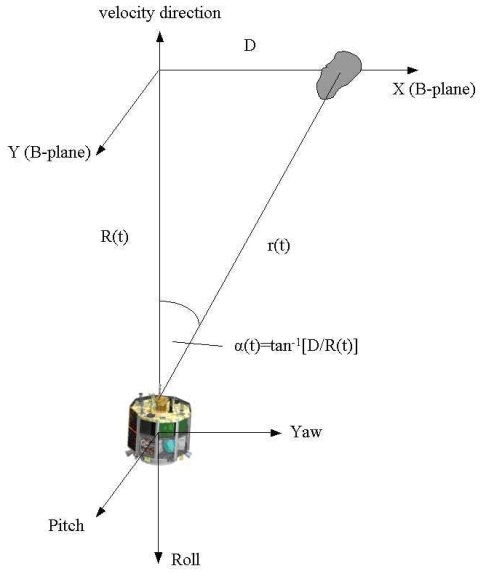


Figure 2: NEO Encounter Geometry

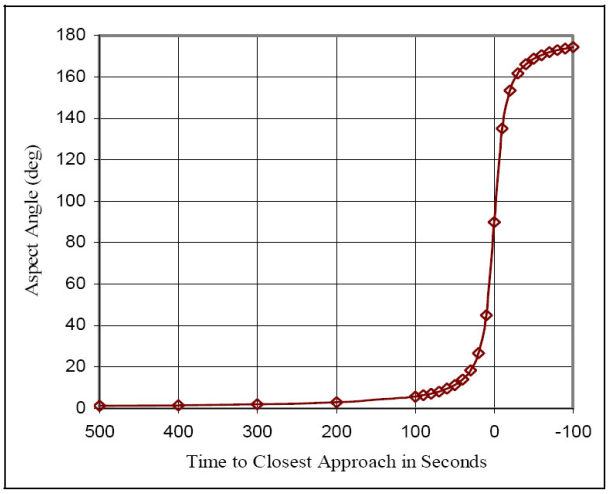


Figure 3: Toutatis Aspect Angle: $\alpha(t)$

approach. The expression for the rate of change of the pitch angle $\alpha(t)$ follows from eq. 1:

$$\frac{d\alpha}{dt} = \frac{VD}{(D^2 + V^2)} \tag{2}$$

Fig. 4 shows its behavior near the time t_{CA}. On the basis of the adopted values for the flyby velocity V = 10 km/s and the minimum miss distance D_{min} = 100 km, we find that the maximum required pitch rate (at the time of closest approach) equals 5.7 °/sec. This result forms the basis for the sizing of the CMG control capability.

3. Spacecraft Design

The NEO2M Mini-satellite platform will be derived from Surrey's existing spacecraft configurations [1, 2, 3]. Its diameter is 1.4 m and its height is 0.8 m. The thrust tube is made of 2 mm thickness aluminum alloy sheet and is capable of supporting an additional spacecraft mounted on top for a dual launch. The preliminary mass budget for a spacecraft design with a 10-kg science reference payload is provided in Table 1. The satellite moments of inertia are estimated to be $[I_{xx}, I_{yy}, I_{zz}] = [53, 53, 60]$ kg m^2.

4. ADCS Architecture

4.1 Selection of Actuators

The objective of the ADCS is to support the mission during all its phases. Of particular importance is the implementation of the high rotation rate required for the NEO imaging around the time of closest approach. For achieving the required spacecraft pitch rate, there are at least three design options: (i) using a dedicated rotating scan mirror similar to the design implemented for the CONTOUR probe [6]; (ii) using thrusters as actuators; and (iii) employing a Micro-CMG cluster as actuators. Option (i) presents a potential single point failure (without a fall-back option) resulting in loss of mission. Option (ii) would become extremely complex in view of the intricate time dependency of the rate history. Option (iii) has the CMG's tracking the asteroid in a more mass-and power-efficient way than the thrusters could do. Even if one of the gyros fails, a 3-CMG system will still be able to track the NEO, although with a degraded performance. Furthermore, there may still be a third level of redundancy by using a combination of CMG's and thrusters. Therefore, a CMG-based ADCS system will provide an efficient and redundant means of tracking the NEO around close approach. Finally, it may be noted that the use of CMG's offers a substantial amount of flexibility in the case of a mission extension with additional flybys of other candidate NEO's. Following a detailed trade-off of a number of sensor combinations, the baseline ADCS suite shown in Table 1 has been selected. Table 2 presents a summary of the mass characteristics of the NEO2M Minisatellite platform.

5. Attitude Control by Micro-CMGs

The most critical ADCS control objective of the mission occurs around the time of closest approach with the NEO object. The spacecraft must be rotated about its pitch axis with a rate of close to 6 °/sec in order that the imager keeps pointing towards the NEO. The evolution of the pitch angle and rate as a function of time are shown in Figs. 3 and 4 and form the basis for the design of the CMG control law implementation. Micro-CMG's are the baseline actuators for performing the three-axis attitude control because of their low-mass and low-power properties and agility. Surrey Satellite Technology Ltd. (SSTL) has considerable in-flight experience with these CMG's [6] so the risk of using this hardware on a NEO flyby mission is acceptable. Table 3 summarizes the design characteristics of the proposed SSTL Micro-CMG's and Fig. 5 provides a visualization of the CMG elements.

The maximum attitude control torque that needs to be delivered by the CMG's must be determined on the basis of the maximum required spacecraft rotation rate of 6 °/sec, the spacecraft spin axis moment of inertia of about 60 $kg - m^2$ and the interval of time allowed to achieve the maximum rotation rate (taken as 3 sec). The required angular momentum h_0 per CMG for a 4-CMG cluster in a skewed pyramid configuration can be calculated as:

$$h_0 = \frac{1}{2} H/(1 + \cos\beta) \approx 0.317H \tag{3}$$

where H stands for the total angular momentum of the CMG cluster and $\beta = 54.7°$ is the skew angle. Finally, the gimbal rate for the control execution must be calculated from the applied torque and h_0.

From previous SSTL's CMG experience [6] it appears feasible that the required angular momentum of 1.99 Nms can be achieved with a gimbal rate of about 85°/s. In order to achieve the CMG angular momentum h_0, a flywheel with angular speed $\omega_w = 6000$ rpm and a spin moment of inertia of 0.0032 kg m^2 has been selected based on SSTL's experience with reaction/momentum wheels and Micro-CMG's.

For the ADCS systems, if we use a bandwidth of 0.5 Hz (typical for Surrey Satellites) the bandwidth of the gimbal motor dynamics should be at least 10 times higher at 5 Hz (i.e. accelerate to the maximum gimbal rate of 1.5 rad/s in 0.2-0.5s). For a gimbal rate of 1.5 rad/s with a 33.3% performance margin backup (i.e 2 rad/s) the gimbal acceleration to be provided from the gimbal motor will result to an acceleration of 10 rad/s^2. From this and if we assume (preliminary assumption) that the inertia of the gimbal motor will be dominated by the flywheel MOI:

$$N_g = I_g \delta_{max} \tag{4}$$

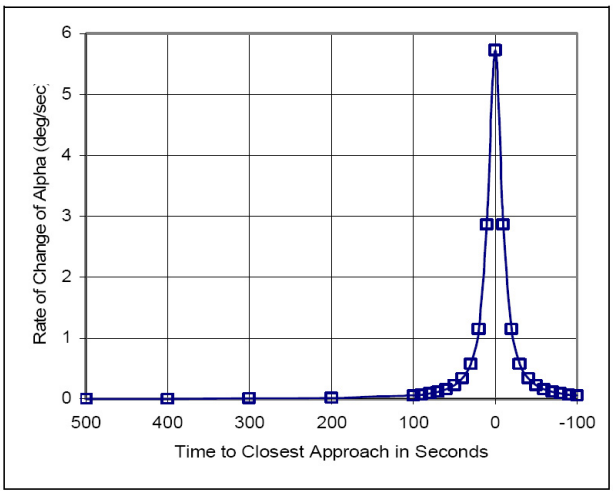


Figure 4: Rate of Toutatis Aspect Angle: $d\alpha/dt$

Table 1: NEO2M Platform Characteristics

Component	Units	Mass (kg)	Power (W)	
			Peak	Idle
Micro-CMG's	4	2.5	10	0.2
V-Slit Sun Sensor	1	1	1	0.1
3-axis Sun Sensor	1	0.3	0.1	0.02
BAE Gyros	4	0.035	0.25	0.05
Star Camera	2	1.7	2.8	0.3
Thrusters	8	0.28	10	0.2
Margin (20%)		3.4	10	0.7
Total		< 20	< 150	< 4

Table 2: ADCS Mass and Power Budgets

Subsystems	Mass (kg)
Payload	10.0
Propulsion	208.5
Structure (includes harness and solar arrays)	54.3
Attitude Determination and Control	19.5
Power	16.2
Communications	11.6
Environment (radiation and thermal)	7.1
On Board Data Handling	2.8
Margin (20 %)	31.3
TOTAL	396

Table 3: Specifications of SSTL Micro-CMG's

Satellite MOI	(53, 53, 60) $kg - m^2$
Maximum Slew Rate ω_{max}	6 °/sec
CMG Cluster Skew Angle β	54.7 °
Total Angular Momentum H	6.28 Nms
CMG Angular Momentum h_0	1.99 Nms
Torque N_{CMG}	2.0944 Nm
Gimbal Rate	1.315 rad/sec
CMG Flywheel Inertia Moment	0.00235 $kg - m^2$
Flywheel Speed ω_w	6000 rpm

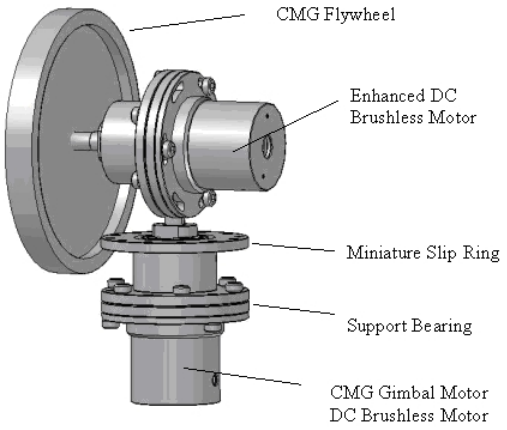


Figure 5: Conceptual Micro-CMG Design

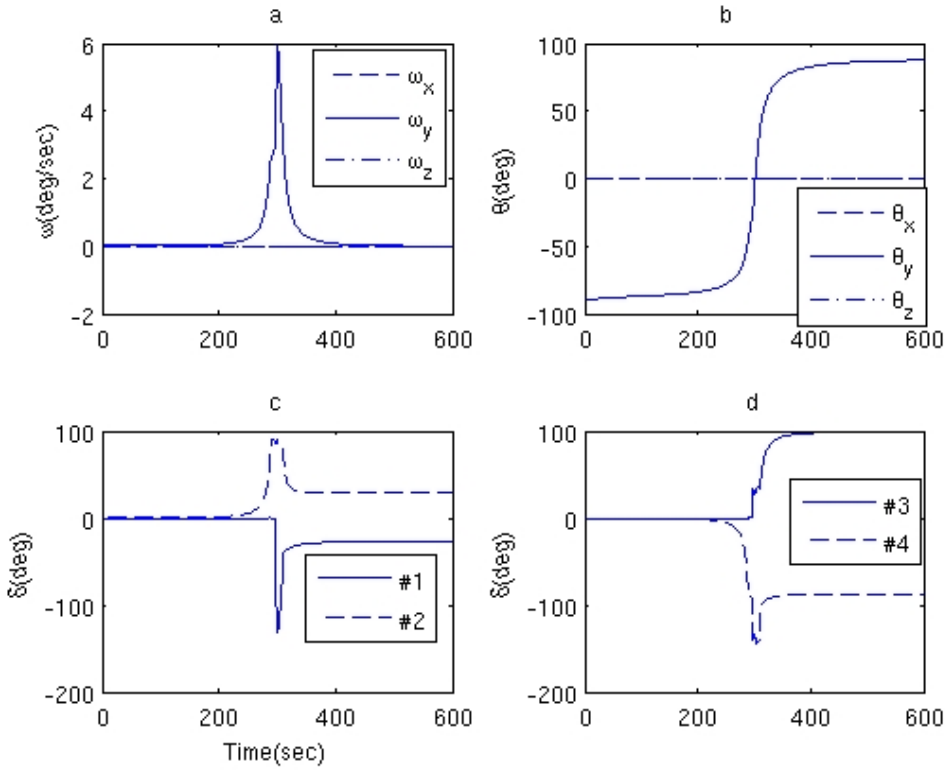


Figure 6: Conceptual Micro-CMG Design

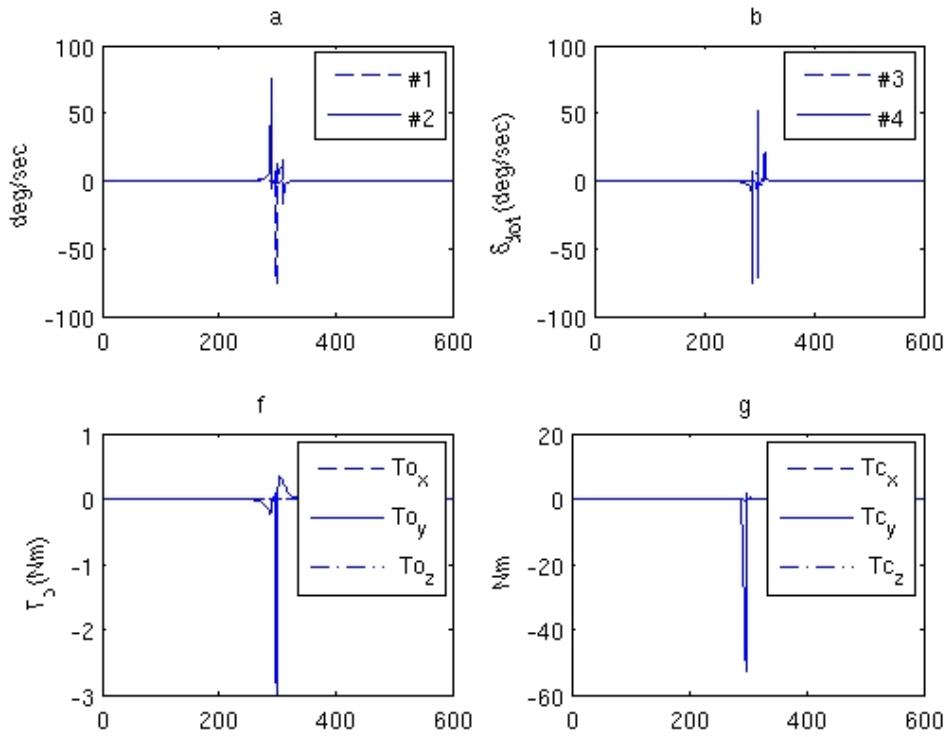


Figure 7: Conceptual Micro-CMG Design

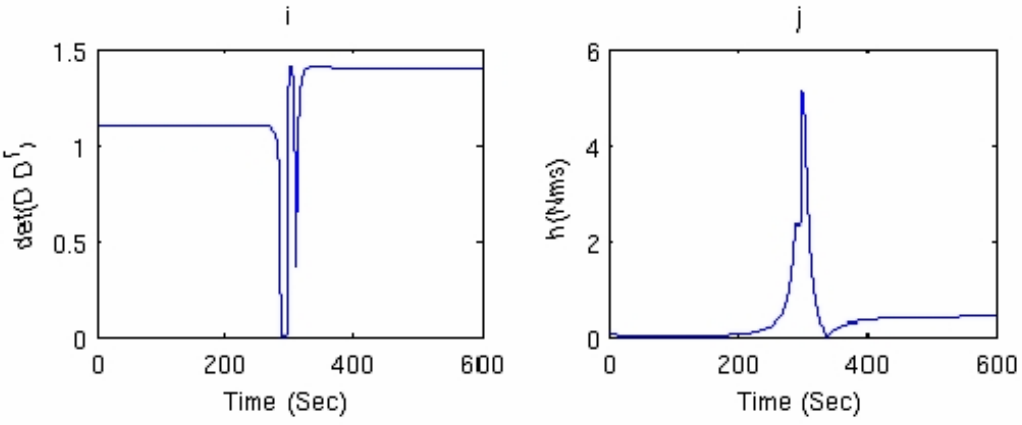


Figure 8: Conceptual Micro-CMG Design

For a gimbal MOI $I_g = 0.004kg - m^2$ and a gimbal acceleration of 10 rad/s^2, $N_g = 40$ mN-m (or $4x10^{-2}$ Nm). Thus the torque amplification factor for the CMG would be:
The torque amplification factor is similar to that of other CMG units available.

5.1 Agile Maneuvers with CMGs

The following simulations use the Generalised Singularity Robust singularity avoidance logic [7, 8]. The CMG sizing specifications have given a design specification for a 4-CMG cluster (Table 3). A simulation for the required NEO tracking maneuver is provided in Figures 6. The PID logic and singularity avoidance scheme used in the NEO tracking maneuver have been presented in detail in Wie [7, 8]. In summary:

$$\boldsymbol{\tau} = -\mathbf{J}\left\{2k\underset{L_i}{sat}\left(\mathbf{e} + \frac{1}{T}\int \mathbf{e}\right) + c\boldsymbol{\omega}\right\}$$

$$L_i = \frac{c}{2k}\min\left\{\sqrt{4a_i\left|e_i\right|}\left|\omega_i\right|_{max}\right\}$$

$$\dot{\boldsymbol{\delta}} = \mathbf{A}^{\#}\dot{\mathbf{h}}$$

$$\mathbf{A}^{\#} = \mathbf{A}^{\mathbf{T}}\left[\mathbf{A}\mathbf{A}^{\mathbf{T}} + \lambda\mathbf{E}\right]^{-\mathbf{1}}$$

$$\mathbf{u} = -\boldsymbol{\tau} - \boldsymbol{\omega} \times \mathbf{h}$$

where,

$$\left|\mathbf{E} = \begin{array}{ccc} 1 & \epsilon_3 & \epsilon_2 \\ \epsilon_3 & 1 & \epsilon_1 \\ \epsilon_2 & \epsilon_1 & 1 \end{array}\right| > 0 \qquad (5)$$

λ is a scalar
ϵ_i is selected as a modulation function: $\epsilon_i = \epsilon_0 \sin(\omega t + \phi_i)$
ϵ_0 is the amplitude
ω is the modulation frequency
ϕ_i is the modulation phase offset
The results show that the proposed CMG's are capable of performing the required maneuver with a comfortable margin and without exceptionally large excursions of the gimbal angle.

6. Conclusions

A preliminary design of an autonomous guidance and control system for a mini-satellite performing a Near Earth Object (NEO) fly-by mission has been proposed. After the transition from spin to three-axis mode, the attitude knowledge is provided by a set of low-cost gyros in combination with a star sensor. For the execution of the tracking control during the critical encounter phase, a Micro-CMG cluster has been selected and its specifications have been determined. The CMG's performance has been assessed by means of realistic simulations. The CMG's prove to be an efficient means of providing the agility required to track the NEO for the imaging. The results demonstrate the practicality and versatility of the proposed guidance and control concept for low-cost NEO deep space missions.

Bibliography

[1] S. Jason, A. da Silva Curiel, L. Gomes, A. Phipps, J. Ward, and W. Sun. Low cost planetary exploration: Surrey lunar minisatellite and interplanetary platform missions. In *IAF-00-Q.4.02*, 2000.

[2] A. Phipps, A. da Silva Curiel, L. Gomes, and J. Ward. The moon and inner planets - getting there for low cost. In *IAA-L-0403*, 2000.

[3] A. Phipps et. al. An entry level mission to a near earth object. In *Proceedings of 5th IAA InternationalConference on Low-Cost Planetary Missions*, number ESA SP-542, pages 241–249, November 2003.

[4] 2002. `http://echo.jpl.nasa.gov/asteroids/4179_Toutatis/toutatis.html` Asteroid Radar Research web page, Jet Propulsion Laboratory.

[5] F. Kennedy et. al. Prometheus: A low-cost micro-satellite flyby mission of 4179 toutatis. In *Proceedings of 5th IAA International Conference on Low-Cost Planetary Missions*, number ESA SP-542, pages 227–238, November 2003.

[6] V. Lappas, P. Ooisthuizen, P. Madle, L. Cowie, G. Yuksel, and D. Fertin. Design, analysis and in-orbit performance of the bilsat-1 microsatellite twin control moment gyroscope experimental cluster. In *AIAA GNC Conference, Providence, Rhode Island*, number Paper AIAA-2004-5246, Aug. 16-19, 2004.

[7] B. Wie, C. Heiberg, and D. Bailey. Singularity robust steering logic for redundant single-gimbal control moment gyros. *Journal of Guidance, Control, and Dynamics*, 24(5):865–872, 2001.

[8] B. Wie. Singularity escape/avoidance steering logic for control moment gyro systems. *Journal of Guidance, Control, and Dynamics*, 28(5):948–956, 2005.

Solar Kites for Earth Magneto-Tail Monitoring

Vaios Lappas[1], Bong Wie[2], Colin McInnes[3], Lorenzo Tarabini[4], Luis Gomes[5], Kotska Wallace[6]

[1] Surrey Space Centre, University of Surrey, GU2 7XH, United Kingdom, E-mail: v.lappas@surrey.ac.uk
[2] Arizona State University, Tempe, AZ 85287-6106, USA, E-mail: bong.wie@asusat.edu
[3] University of Glasgow, Glasgow, G12 8QQ, UK, E-mail: c.mcinnes@glasgow.ac.uk
[4] GMV S.A. C/Isaac Newton, Cantos, 28760 Madrid, Spain, E-mail: l.tarabini@gmv.com
[5] Surrey Satellite Technology Ltd, Guildford, Surrey, GU2 7NE, UK, E-mail: l.gomes@sstl.co.uk
[6] ESA, ESTEC, Noordwijk, The Netherlands, E-mail: K.Wallace@surrey.ac.uk

Abstract

Solar Sails have been studied in the past as an alternative means of propulsion for spacecraft. Recent advances in Solar Sail technology and the miniaturisation of technology can drive these systems much smaller ($<$ 5 kg mass, $<$ 10m sail diameter) than existing sails, while still having a high delta-V and acceleration capability. With these unique capabilities of miniature Solar Sails, called Solar Kites, some very unique space science missions can be achieved which are difficult to be implemented using conventional propulsion techniques. One such unique candidate mission is to study the Earth's magnetotail. The paper lays out the main design features and technologies of a Solar Kite mission/platform and demonstrates that a cluster of Solar Kites with science payloads can provide multiple, in-situ measurements of the dynamic evolution of energetic particle distributions of the rotating geomagnetic tail of Earth. With a unique design, a Solar Kite proves to be an efficient, affordable and versatile solution for the mission analysed with a significant science return.

1. Proposed Mission

The geomagnetic tail around Earth poses an important scientific problem related to weather conditions on Earth. There are a number of missions studied to date which focus on using a high number of nanosatellites (up to 100) for continuous multipoint measuring of the field. Conventional geomagnetic tail missions require a spacecraft to be injected into a long elliptical orbit to explore the length of the geomagnetic tail. However, since the orbit is inertially fixed, and the geomagnetic tail points along the Sun-Earth line, the apse line of the orbit is precisely aligned with the geomagnetic tail only once every year. Approximately 4 months of data can be acquired, with only 1 month of accurate data from the tail axis. To artificially precess the apse line of the elliptical orbit to keep the spacecraft in the geomagnetic tail during the entire year would be prohibitive using chemical propulsion. A scientifically interesting 11 x 23 Earth radii elliptical orbit would require a ΔV of the order 3.5 km/s per year of operation for apse line rotation. A perigee at 11 Earth radii meets the bow shock, while an apogee at 23 Earth radii is optimum to observe magnetic reconnection in-situ. Although the ΔV for apse line rotation is large, artificially precess the apse-line of 11 x 23 Earth radii orbit, stationing a miniature science payload permanently within the geomagnetic tail and so providing continuous science returns. Using multiple solar kites ($\sim$ 35), the entire geomagnetic tail could be populated by sensors which precess with the annual rotation of the geomagnetic tail, allowing real-time visualisation of the 3D plasma structure of the geomagnetic tail. Such a real-time visualisation would provide insight into the fundamental plasma physics of the geomagnetic tail [1].

2. Mission Analysis

In order to showcase the potential advantages of SKs and sailcraft in general, a comparison is presented of sailcraft and specifically a SK with other propulsion techniques for the Geosail mission. To perform the comparison with different types of propulsion, two types are used: Solar Electric Propulsion (SEP), Chemical propulsion,

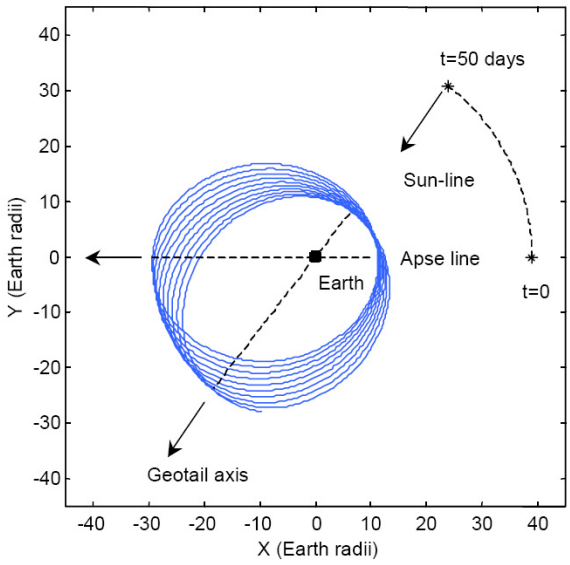


Figure 1: Evolution of Elliptical Orbit at ∼1 deg/day

represented by a generic motor. The main inputs for the analysis are the total required ΔV and the acceleration. Since the bus platform mass is limited to 1.5kg (as this is the mass of the SK concept), and the characteristics of the sail and associated structure are pre-defined, the characteristics of the required sail can be obtained. The targets are presented in Table 1.

To reach the targets of the mission, the solar sail propulsion system would have to have the characteristics of Table 2. For the comparison with chemical and SEP, some assumptions need to be made. I_{sp} is the specific impulse for the propulsion system used (typical values of 300 s for chemical propulsion and 3300 s for SEP are used throughout). If the assumption of a thrust to mass ratio of 625 is made, or for example, a 1 mN thruster is used, then the mass of the propulsion system will be 1.6 kg. In the case of chemical propulsion, it was estimated that for a low propulsion system total mass (propulsion system + propellant) the propellant mass is 2/3 of the total mass. For higher propulsion system total masses ($>$ 20 kg) it is assumed that the propellant mass is 90% of the total mass. This heuristic approach is based on current and future SSTL missions. For an SEP system:

The thrust selected for this engine is of 1mN, much higher than required for the mission acceleration, but it is not desirable to go below this value. The above assumptions assume that SEP technology is available for such small thrust, mass and power levels. The use of a chemical system is not realistic in this case as the total mass of the mission is likely to reach something in the order of 22 kg. In Figure 2 the comparison between the extra mass required by the solar sail and the SEP (not require contingency) is presented. The mass advantage of a mission based on the SK concept is evident, and in this case, since the duration of the mission is fixed there is no disadvantage due to the long time spans involved.

2.1. Payloads

The SK has a space science objective, to study the geomagnetic tail, allowing phenomena such as magnetic reconnection to be studied in-situ. The goal is to use a constellation of multiple solar kites (∼35), thus the entire geomagnetic tail could be populated by sensors which precess with the annual rotation of the geomagnetic tail, allowing real-time visualisation of the 3D plasma structure of the geomagnetic tail. Such a real-time visualisation would provide insight into the fundamental plasma physics of the geomagnetic tail. The large number of SK can allow using different configurations of payload suites. For example if a constellation (∼35) of SK's is used to study the earth's magnetic tail, most of them can carry magnetometers and plasma detectors but a small number can carry space dust detectors to complement and maximise the science return from the mission. The suggested ultra miniature payloads are presented in Table 4.

3. Solar Kite Design

Reviewing existing solar sail technologies it is determined that a realistic assumption to begin designing a sailcraft is to use a Sail Assembly Loading (SAL) factor of 10 g/m^2. This value depends on the availability of solar sail technology. The SAL is defined as:

$$SAL = \frac{Mass of Sail Structure}{Solar Sail Area} = \frac{m_s}{A} \qquad (1)$$

Using Eq. 1 and some initial condition values such as the dimensions of the solar sail, one can deduce the design

Table 1: Targets for the GeoSail Mission

SK Desired Characteristics	
Total ΔV (km/s)	3.5
Acceleration (m/s^2)	1.11E-04
Bus and P/L Mass (kg)	1.5

Table 2: Solar sail performance

SK Sizing Parameters	
Solar Sail Mass (Ms)	0.235 kg
Total Mass	1.735 kg
Total Mass (+margin)	2.256 kg
Sail Area	23.814 m^2
Sail Side	4.88 m
Sail Film Mass (m$_f$)	0.071 kg
Mass of Booms	0.137 kg
Mass of Mechanisms	0.027 kg
Sail Structure Mass (m$_b$)	0.164 kg

Table 3: SEP Performance

SEP Platform	
Mass of Propellant	0.354 kg
Mass of Motor	1.6 kg
Mass of Solar Panels	N/A
Total Mass	3.454 kg
Total Mass (+margin)	4.490 kg

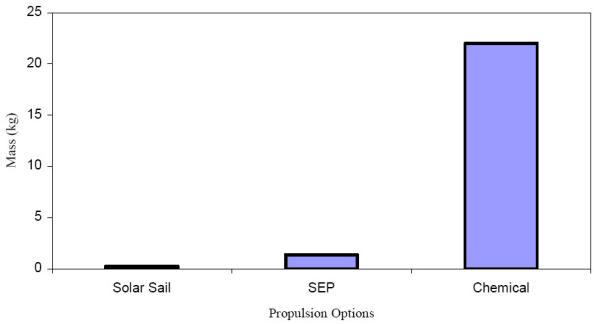


Figure 2: Solar Kite Propellant Mass Required

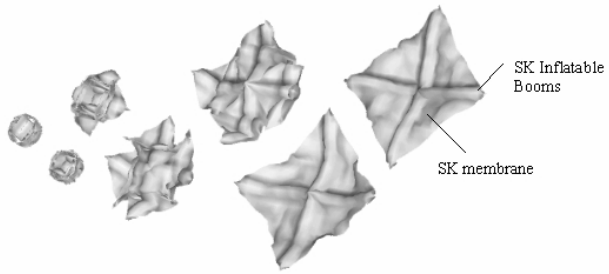


Figure 3: SK Booms, Sail and Deployment (Nihon Concept)

Table 4: Solar Kite Candidate Payloads

Payload	Science	Mass (kg)	Power (W)
Mag/ter	Magnetic Fields	$<$ 100 g	$<$ 0.5 W
PIE Detector	Proton/Ion Detection	$\sim$ 400g	$<$ 2W
Plasma Detector	Plasma Bubbles	$<$ 300g	$<$ 2W
E. Sensor	Space dust	$<$ 50g	$<$0.2W

parameters of the SK: For a square sail of 5 m $\times$ 5 m, A = 25 m^2 and using Eq. 1 and the given SAL the mass of the sail structure m$_s$ becomes m_s = 0.25 kg. The boom-sail structure has a 4 cm diameter and a length of L = 3.535 m. The mass of the film used is m_f = 0.05 kg. The mass of the SK platform including the platform subsystems and payloads is assumed to have a mass of m_bus = 1.5 kg or less. The mass of the bus, solar sail area and SAL are noted to be key in the design of the SK. Using the values above it is then deduced: For a SRP constant of $P = 4.536 \times 10^{-6} N/m^2$ and a thrust coefficient η = 1.8, the maximum thrust of the SK is $F_{max} = \eta PA = 2.04 \times 10^{-4} N$. Then the area-to-mass ratio $r_{a/m}$ is 10 m^2/kg and the areal density $\sigma = m/A$ = 0.1 kg/m^2. The acceleration is then:

$$a_c = \frac{F_{max}}{m} = \frac{\eta PA}{m} = \frac{\eta P}{\sigma} = 1.2 \times 10^{-4} ms^{-2} \quad (2)$$

The value for the acceleration of the SK is comparable to those for other sail missions currently under design. All values derived are summarised in Table 5. The acceleration calculated is able to produce the required acceleration of $1.1 \times 10^{-4} m/s^2$.

3.1 Solar Sail Technology

The analysis of existing and future developments on solar sail boom technology has led to a number of important conclusions:

1. The mass per length ratio for the booms is a critical, mission enabling factor
2. Conventional and current boom technologies can't be scaled down to an SK scale (3.535 m boom) and preliminary analysis indicated that this technology has a use threshold for solar sails of $>$ 20 m
3. Sails of $<$ 20 m will require mass per length ratio (specific mass) $<$ 60 g/m
4. Deployment of solar sail booms is complex and has been analysed for large ($>$ 40 m sails), making this technology difficult to implement on a SK
5. SK will need a simple, ultra light sail with a smaller life time from large sails

A SK with a 3.535 m boom will need to be a simple and optimised design to a 1.75 kg spacecraft mass. The small size of the boom can prove instrumental in this in that a simpler, less complex boom can be manufactured and deployed compared to existing 40 m booms with multiple motors, pulleys, supports that add risks to the sail design, mass overheads and complexity [2]. A semi-active deployed boom is proposed consisting of rigidized inflatable material with an integrated sail to the booms and with a simultaneous boom and sail deployment. This integrated approach brings significant mass/volume savings as well as a simple deployment strategy for the SK sail.

Sail film and supporting structure technologies play a key role for the realisation of solar sail design concepts. Ultra-thin film of the order of 1-2 micrometer of polyimid basis have already been manufactured under laboratory conditions. This is an advantage for SK, since producing limited quantities in lab conditions is sufficient. The SK sail membrane is a 0.9 micrometer polymid based film, which is based on the DuPont polymid membrane. A similar version (1 μm) is used in the L'Garde solar sail mission [3].

The main design requirements are the 5 m length of the SK sail, compact packaging, simplicity and robustness of deployment and a $<$ 60 g/m mass per length factor. The most optimum material able to achieve this are rigidized inflatable structures. The SK team has chosen to use an integrated approach for the SK boom-sail film-deployment design. In this design the sail film/membrane is integrated with the booms. Deployment is achieved using an inflating gas. This integrated semi-active approach is able to bring significant savings in mass, volume and power to the SK, not requiring motors, extra electronics, pulleys or complicated mechanical structures for boom deployment [4]. The two biggest advantages of using the inflatable rigidized boom/sail are the high density packaging capability and the $<$ 60 g/m mass per length factor. The SK rigidized boom is blended with the 1 μm sail film. A 4 cm diameter boom is able to provide the necessary structural rigidity of 200 MPa pressure needed to sustain various loads in space, as analyzed in various inflatable structures currently in design, including margins. Many institutions are working on the development of rigidisable inflatable structures with promising results. Nihon University has conducted experimental tests completed in microgravity conditions of a 1 m inflatable boom [5]. The goal in the Nihon experiments is to demonstrate inflatable technology using a 1 kg cubesat The only shortcoming of this technology

Table 5: Solar Kite Characteristics

Solar Kite Parameters	Values
Sail film + Booms +Depl. Mech.	0.2 kg
Length of Booms *L*	3.535 m
Bus/Payload m_{bus}	1.5 kg
Total mass $m_{s/c}$	1.75 kg
Sail Area *A*	25 m^2
Thrust Coefficient η	1.8
Acceleration a_c	0.12 mms^{-2}

is the need to completely study the phenomenon of wrinkles, an issue still researched for conventional sails. Deployment is achieved by two miniature valves, identical to the propulsion valves used in the SK ADCS system. A 9 g gas will inflate the structure and be able to provide continuous pressure for a minimum 2 year lifetime of the SK. The calculated volume for the SK boom/sail structure is the smallest possible since storage for the integrated 'structure' is much more compact and lighter than using a traditional CFRP design.

The technology readiness of the suggested technology, is at TRL 4 (i.e flown in space but in need of modification, customisation or optimisation for specific application) for inflatable structures and 5 for rigidized structures. Using the suggested sail-boom-deployment concept with the specified parameters above (4 cm boom diameter, 3.535 m length) the mass breakdown using the three available technologies (CFRP, coilable, inflatable) is depicted in Table 6. The inflatable option provides significant savings in mass and volume on the overall SK design. A CFRP option is too large for a SK mainly due to its large mass/boom meter ratio and deployment mechanism. A coilable option is close to the 2.5 kg mass requirement though it comes with a high level of complexity in deployment. Minimum storage and an efficient ultra low mass/length ratio makes the inflatable option in its integrated design a mission enabling technology, with most of its technology available or tested.

3.2 Sail Deployment

The sail-boom integrated structure is deployed with a gas based inflation system. The inflation system is a continuous inflation system consisting of two simple gas valves slightly modified from the ultra miniature resistojet thruster used in the SK ADCS system. The system contains two valves, one per boom (two booms). A small gas tank in a ring configuration is used split in two parts, one side containing Helium and the other sealed side containing low-pressure liquid hydrazine (LHZ). The Helium gas is initially used for inflation of the sail/boom structure and then LHZ is used as a 'make-up gas' to continuously keep the inflated structure rigid. Hardening strips are also used with a special curing coating to assist a fast curing process when the SK sail is deployed and points to the sun. A COTS canister has been proposed for similar applications and has been proposed to be 2.6 cm long, 0.8 cm in diameter, and have a volume of 8.2 cm^3, if using a Helium gas. The helium in the canister will be stored at 60.5 psi. Once the helium is released from the canister at 0.15 cm^3 per second, it will take five minutes for the canister to extinguish the helium supply. This will leave a final pressure in the canister of 0.5 psi. The SK will have a two-year lifetime and the LHZ required maintaining two inflatable booms (two diagonal) is calculated to be 15 g. With two valves and a spiral miniature tank made out of aluminum, the system will weigh 90 g. With such a small mass the deployment mechanism is to small to consider for ejection (if it was not necessary for use) and besides its task to inflate-deploy the sail/booms with helium it is needed to maintain the booms and sail rigid and deployed throughout its lifetime. Two pressure valves will be used to measure the gas pressure in the two SK booms and since related to the rigidity of the sail will be used to feed small amounts of LHZ when needed. Hardening strips are a new technology that will be used to assist making the SK a permanently rigid square sail after the sail has been deployed. It is necessary to make sure that the sail remains as rigid as possible in order to have maximum performance. The hardening strips to be used will consist of a tape-like substance that has the unique property that the tape will remain pliable and tape-like until it is exposed to solar radiation. When the strips are exposed to solar radiation, they will begin to harden and will permanently cure in approximately 15 minutes. The hardening strips will be placed on the solar sail in a spider web pattern and along the outer edges to minimize any warping or shape changing of the sail after deployment.

4. Solar Kite Platform

The SK consists of two parts: the SK structure and the SK platform (Figure 4). The SK platform uses as a basis the Surrey Palmsat 1 kg platform. However, the platform is tailored to the SK requirements and mission requirements. The 'bus' is a 6 x 9.5 cm hexagon structure made out of carbon-fibre and aluminum honeycomb structure machined carefully to save mass. The dimensions of the sail are 5 x 5 m with 3.535m inflatable booms integrated with the sail membrane in order to use a single deployment mechanism to save mass again. The thrusters

Table 6: SK Sail/Boom Technologies

SK Sail/Boom Parameters	CFRP (kg)	Coilable (kg)	Inflatable (kg)
Sail film m_f	0.05	0.05	0.05
Booms (4) m_b	0.34	0.25	0.10
Deployment Mechanism	0.1	0.1	0.1
Total	0.49	0.4	0.25

are mounted as such as to be used both in SK sail-boom stored and deployed configuration for SK stabilisation, commissioning and operations. The SK uses three thrusters in a 1 pair plus single thruster configuration. The pair of thrusters are mounted opposite and anti-parallel to each other for spin-up and spin maintenance operations and the third thruster is used for precession.

The thrusters use miniature boards with a battery, solar array (1 x 1 cm) and a communication system based on Bluetooth, based on a distributed wireless link. The SK after being ejected from the launcher or transfer vehicle via SMA bolts and pre-tensioned springs, will switch on its power subsystem and de-tumble using the ADCS thrusters. Having established an RF link with the ground station, the SK will start to deploy its integrated sail-boom structure in less than seven minutes using a stored helium gas. Figure 5 shows the conceptual SK design.

Figure 6 details the internal view of the SK sail-boom compartment. The inflation gas ring tank is split into two parts one for the helium inflation gas and the other for the 'make up' LHZ gas. Figure 7 details the SK ADCS thrusters when the SK sail-boom is stowed. The thrusters are mechanically fixed in their 'stowed' position and when the sail-boom inflates the tips of the sail-boom structure are freed from that position and move freely.

5. Launch and Deployment

Surrey has worked on in-house transfer vehicles for such missions as the SK. A modified version of the Surrey Transfer Vehicle (STV) designed for the SK mission, will have a single 400 Newton, N_2O-HDPE Hybrid system based on the Daimler Benz S400/2. The STV-SK vehicle will have as a main requirement to take the SKs from a 580 x 35786 km and 7° inclination to the desired [(11x23) x Re] GeoSail orbit. The STV-SK vehicle will have to fit to an Ariane 5 ASAP mini-satellite space, which has a 150 cm diameter and 150 cm height, and the STV will need to achieve a 1400 m/s ΔV. For the selected Daimler Benz S400/2 engine with an I_{sp} of 318 s, translates to 109 kg fuel. After some system design and analysis this leaves 80 kg for SKs ($\sim$35 SKs) and 111 kg for the STV-SK structure and subsystems. Surrey is currently analysing such a system to transfer SK's to the GeoSail orbit as part of a study to design a small satellite space weather-monitoring constellation for ESA.

6. Conclusions

Most of the studies done to date assume that the largest obstacle in solar sail (SS) missions is the required development of the necessary SS specific technologies such as membranes, large stiff and light booms and pointing mechanisms, which is partly true. One of the enabling factors though that make SS missions possible is the miniaturisation of the spacecraft bus, bringing the overall spacecraft mass down and thus enabling solar sails to materialize. A SK with a simple and robust design, equipped with niche scientific payloads can be a significant tool to space planners. It has become clear in the analysis designing a SK mission, that solar kites can provide a number of key advantages when compared to larger, more complicated and expensive solar sails:

- Complexity of solar sails can be reduced
- Sail deployment can potentially be simplified
- Ultra miniaturisation and MEMS technology increase sail acceleration
- Small size sails can use inflatable technology (ultra light) for sails $<$ 7-10m
- Miniaturisation technology is near available
- Smaller spacecraft design times, multiple spacecraft available-constellations

Bibliography

[1] C. R. McInnes, M. Macdonald, V. Angelopoulos, and D. Alexander. Geosail: Exploring the geomagnetic tail using a small solar sail. *Journal of Spacecraft and Rockets*, 38(4):622–629, 2001.

[2] M. Leipold, C.E. Garner, R. Freeland, A. Herrmann, M. Noca, G. Pagel, W. Seboldt, G. Sprague, and W. Unckenbold. Solar sail technology development and demonstration. In *4th IAA Int. Conference on Low-Cost Planetary Missions*, 2000.

[3] D. Cohen, P. Gloyer, and J. Rogan. Preliminary design of a high performance solar sailing mission. In *16th Annual USU/AIAA Conference on Small Satellites, SSC02-II5, Logan, Utah2002*, 2002. accessed January 23rd 2004, `http://aria.seas.wustl.edu/SSC02/papers/ii-5.pdf`.

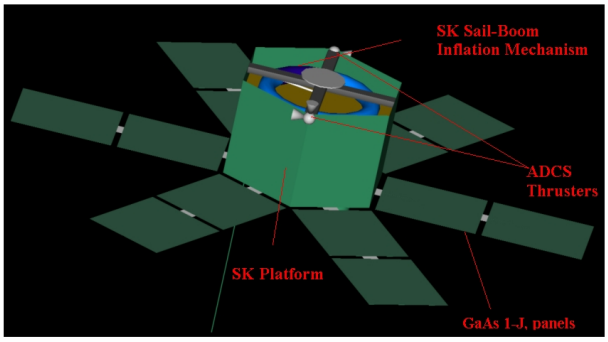


Figure 4: SK Platform and Sail-Booms Stored

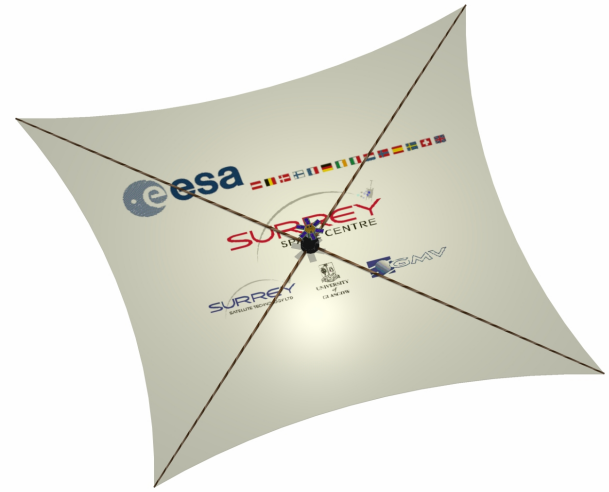


Figure 5: Solar Kite Conceptual Design

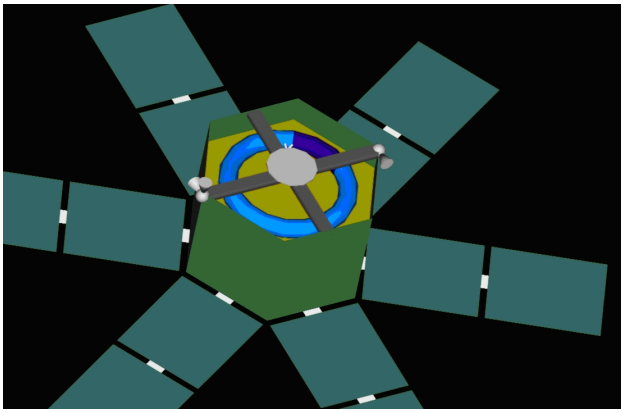


Figure 6: Internal View of the SK Sail-Boom Compartment (Membrane is Hidden)

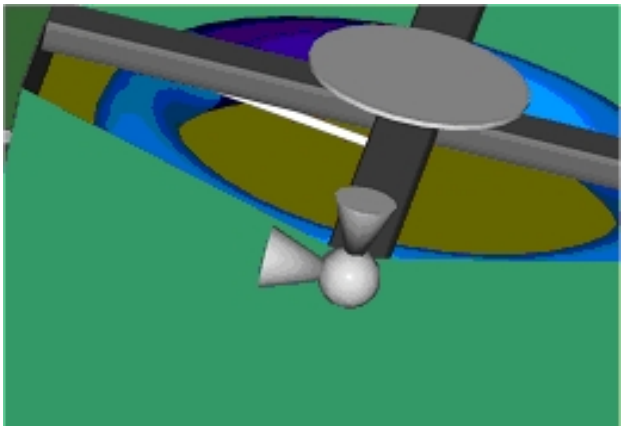


Figure 7: SK ADCS Thrusters (Membrane is Hidden)

Table 7: Solar Kite Mass and Power Breakdown

Subsystem	Mass	Power
Payload	500 g	1.5 W
OBC	150 g	1.5 W
RF System	280 g	5 W
ADCS	220 g	1.1 W
Power	150 g	0.5 W
SK Sail/Boom	200	1.0 W
SK Structure	250 g	-
TOTAL	1750 g	10.6 W
Margin (30%)	525 g	3.18 W
SK Total	2275 g	13.78 W

[4] D. Cadogan. *Gossamer spacecraft: Membrane and inflatable structures technology for space applications*, chapter Rigidization Mechanisms and Materials, pages 257–279. AIAA, Reston. Va. 191, 2001. ed. Christopher H. M. Jenkins and P. Zarchan.

[5] Y. Miyazaki, H. Isobe, T. Kodama, M. Uchiki, and S. Hinuma. Nihon university cubesat program. In *15th Annual AIAA/USU Conference on Small Satellites, Logan, Utah*, October 2001.

[6] C.R. McInnes. *Solar Sailing: Technology, Dynamics and Mission Applications*. Springer Praxis Publishing, 1999.

[7] B. Wie. Dynamic modeling and attitude control of solar sail spacecraft: Part i. In *AIAA Guidance, Navigation, and Control Conference, Monterey, CA, August 5-8,*, number 2002-4572. AIAA, 2002.

[8] D. Gibbon, P. Charman, and N. Kay. The design, development and in-orbit performance of a propulsion system for the snap-1 nanosatellite. In *Proc. 3rd International Conf. on Spacecraft Propulsion, Cannes*, number ESA SP-465, October 2000.

Panel Extension Satellite (PETSAT) - A Novel Satellite Concept Consisting of Modular, Functional and Plug-in Panels

Shinichi Nakasuka[1], Yoshiki Sugawara[1], Hironori Sahara[1], Kanichi Koyama[2], Takanori Okada[2] and Chisato Kobayashi[2]

[1] Department of Aeronautics and Astronautics, University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo, JAPAN, {nakasuka,sugawara,sahara}@space.t.u-tokyo.ac.jp
[2] SOHLA,Creation Core 2102, Aramoto-kita 50-5, Higashi-Osaka, JAPAN, {koyama,okada,chisato}@sohla.com

Abstract

A novel concept of satellite design, named PETSAT, is proposed in this paper. In this concept, a satellite is made of several Functional Panels such as the CPU panel, Battery panel, Communication panel, Attitude control panel, or Thruster panel, each of which has a special dedicated function. By connecting these panels with a reliable connection mechanism in a plug-in fashion, the total integrated system has a full satellite's function. Various combinations of functional panels, (for example, two CPU panels + one communication panels + three attitude control panels + two battery panels, etc.) provide flexibility to deal with various mission requirements, even though the basic panels are the same for various missions. The concept, technical issues and conceptual study results of PETSAT will be discussed.

1. Concept of PETSAT

A novel concept of satellite design, named PETSAT, is proposed in this paper. In this concept, a satellite is made of several Functional Panels each of which has a special dedicated function. By connecting these panels by reliable connection mechanism in plug-in fashion, the total integrated system has a full satellite's function. Various combinations of different kinds of functional panels, sometimes with multiple instances of identical panels, provide flexibility to deal with various mission requirements, even though the basic panels are the same for various missions. These panels are stowed during launch into a small volume (left figure in Fig. 1), and are extended on orbit (right figure in Fig. 1), potentially realizing a satellite with a large antenna, large solar cell area or large boom.
PETSAT intends to change the satellite development cycle in the following way:

1. Each functional panel can be produced in mass quantities so that the reliability can be improved, and the produced panels can be stocked.
2. When a certain mission is given, the satellite bus suitable to the mission requirements can be configured only by connecting the appropriate panels in appropriate quantities in plug-in fashion without much effort needed in ground test of the total system.

This semi-customizable satellite production process is expected to dramatically reduce required manufacture and test time, as well as workload, resulting in a drastic reduction of the individual satellite cost and development time. Mass production of the panels is the key to reducing cost and improving reliability, but in the conventional satellite concept, mass production is difficult even at subsystem level due to the wide variety of mission requirements. PETSAT tries to make this possible by modularizing the various basic subsystem functions into standard panels, and by dealing with the variety of mission requirements with the quantity of the utilized panels of different functions. The concept of PETSAT was initially proposed at the Satellite Design Contest in 2002, where the concept was the winner of the Best Design Award. Five year research project of PETSAT just started in

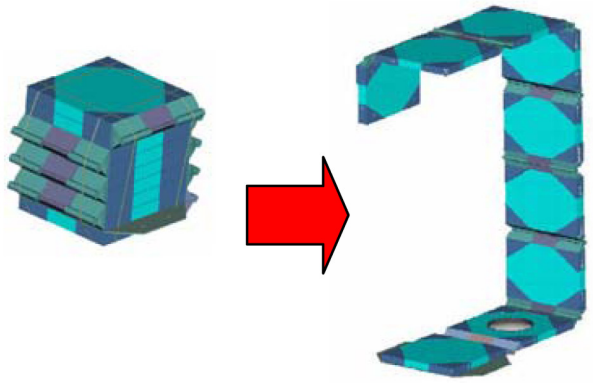


Figure 1: Concept of PETSAT before deployment (left) and after (right)

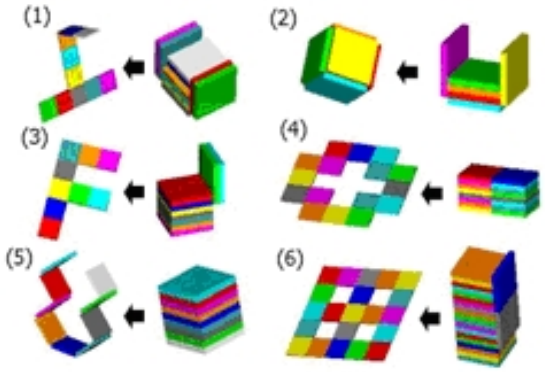


Figure 2: Various shapes realisable by PETSAT

2003, supported by NEDO subsidy, in collaboration of University of Tokyo, Osaka University and manufacturers' community of Higashi Osaka. In addition to the basic concept development and panel and interface designs, we are pursuing the application of the PETSAT concept to nano-scale thunder observation and remote sensing satellites In the following sections, first, some technical issues to be solved to realize PETSAT will be described, followed by the results of conceptual study. In particular, the requirements on the panel internal structure and panel interfaces will be discussed in detail, and the first design results to satisfy the requirements will be given.

2. Technical Issues

2.1 Overall Technical Issues

The key technical issues to realize PETSAT include:

1. The question of how to modularize the satellite bus functions into different panels: for example, whether battery and solar panels should be implemented in a separate functional panel such as an electric panel or should be equipped in every panel as standard basic function, etc.

2. Design of a standard panel structure which provides a suitable thermal and structural environment to different kinds of functional panels since for mass production, different types of panels should use the same structure.

3. Interface between panels: four kinds of interface, mechanical, electrical, information, and thermal interfaces should be carefully designed so that the desired plug-in simplicity of PETSAT can be achieved.

2.2 Requirements on Panel Interface

As for the interfaces, the following requirements should be satisfied.

1. Mechanical Interface
   A panel connection and deployment/latch mechanism, which is very reliable, fault tolerant, soft and highly accurate is required. Also it is required that the panels can be easily plugged-in.

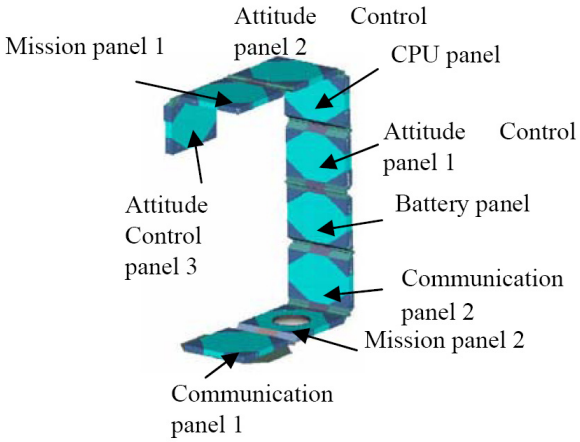


Figure 3: Remote sensing PETSAT

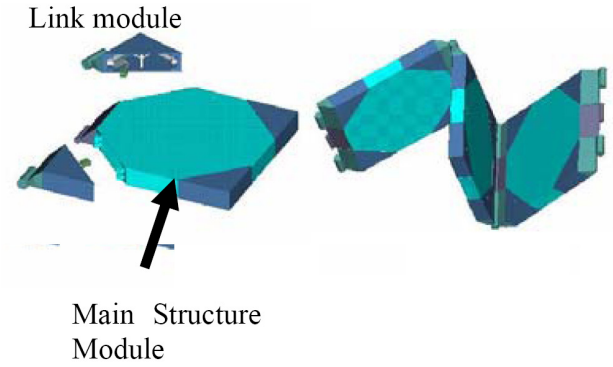


Figure 4: Panel main structure and link module

The accuracy of the angles between panels after being latched should be of a sufficient level to allow the initially planned satellite shape to be achieved. Also the latch mechanism should be soft, which means that some mechanism is required to brake the panels' deployment just before they are latched. Finally, in order to achieve various satellite shapes (see Fig. 2), the sequence of the panel deployment should be carefully designed, otherwise the deployment becomes stacked. Some mechanism is required which assures that, for example, the deployment of panel A and B can be made after panel C is deployed from panel D.

2. Electrical (power) Interface
   In principle, the electric power required in each panel should be generated by solar cells of the same panel, but in many cases power should be transferred from one panel to another panel. PETSAT should have the capability to autonomously transfer power in relation to the power generation and consumption in each panel. Reliability of power lines is another important issue, which should be realized by carefully designed redundancy.

3. Information Interface
   Communication between panels is critical for PETSAT. The information line should be very reliable and should have enough communication capacity to deal with the flow of house keeping data as well as mission data. Each panel should have a microprocessor able to control both this information traffic and manage the information flow inside the panel. So, the total system becomes a multi-processor system where the architecture to manage such large number of CPUs should be carefully designed so that the strengths of a distributed system, such as fault tolerance or capability of grid-computing, can be pursued as much as possible. The information line can be either a wired bus line or an RF line.

4. Thermal Interface
   Thermal coupling within and between panels should be made very strong so that the temperature difference between each part of satellite is as small as possible. [1] indicates that this is the best general strategy for thermal control because (1) the thermal environment is different from mission to mission, and (2) there is not much design freedom

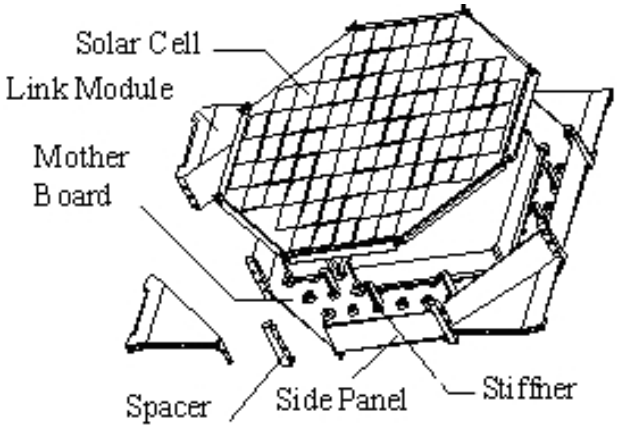


Figure 5: Design of panel main structure

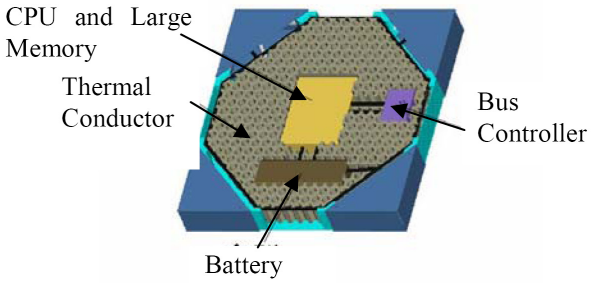


Figure 6: OBC panel inside design

for PETSAT thermal control because the surfaces of panels are almost completely covered with solar cells.

3. Example Missions

Before going into technical details, Let us briefly give descriptions of two example missions realizable by PETSAT.

3.1 Remote Sensing

One potential application is remote sensing. For obtaining high resolution images, the optical system should have long focal length, which requires several mirrors for reflections or a certain distance between a lens and the imaging device such as a CMOS or CCD. Usually several mirrors are utilized for this objective, but high accuracy is required for mirror surfaces, which results in extremely high cost, often with regard to required structural strength. The latter requirement is tough for micro/nano satellites because of size limitations. PETSAT can solve the latter problem through its panel extension mechanism, such as in Fig. 3. Extremely high resolution, such as less than 10 m, is difficult, but a certain level of resolution such as 20 m - 50 m is expected to be achievable with very low cost. Three axis stabilization is required, which is achieved by three attitude control panels, each dealing with stabilization for one axis. In order to obtain well focused images, the accuracy of the angle between panels after deployment is essentially important. However this is very hard to achieve because of the back-rush of the latch or distortion of the panel structure due to temperature change, etc. One method to solve this problem is to control the position and attitude of the image receptor such as CCD or CMOS using very precise actuator. This mechanism is now under development.

3.2 Interferometric Positioning System

If an RF wave arrives at antennae located at different positions on PETSAT, then the difference in arrival time, or phase difference, provides information as to the direction of the RF signal source from the PETSAT. If there are three antennae which are not located collinearly, then the direction of the signal source in 3D space can be estimated. The accuracy of the direction measurement depends on the distance between the antennae (base line) and the knowledge of the relative positions of these

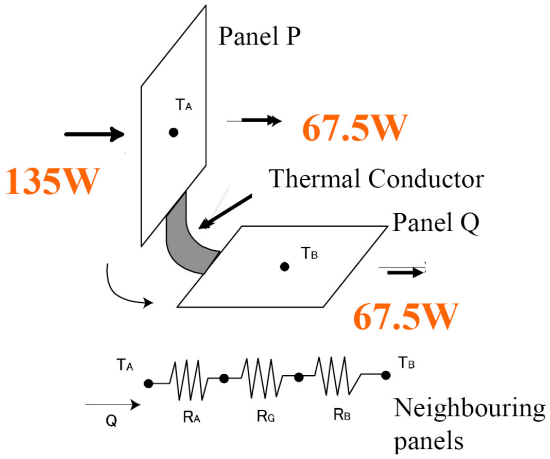


Figure 7: Severest thermal condition for 2 panels

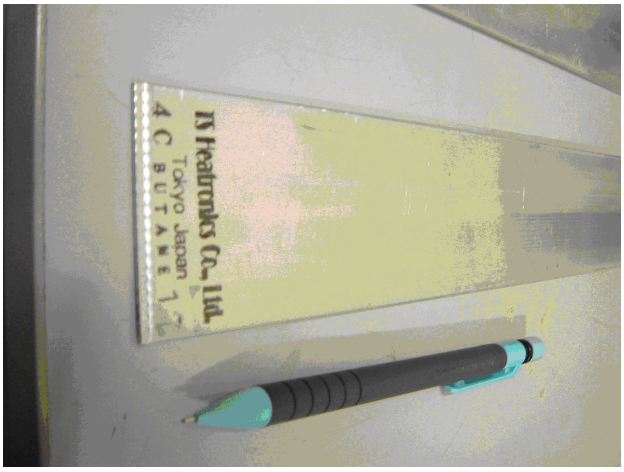


Figure 8: Heat lane plate

antennae with respect to the satellite body frame. Long base line can be achieved in PETSAT by extension of several panels between the antennae panels, and the relative positions of antennae can be estimated by calibrating this sensor system using a signal generated from a known point on PETSAT.

If we want to obtain the location of the RF source on the Earth, then the PETSAT should know its attitude precisely. Several types of navigation sensors for attitude can be employed, including IRU (gyros), magnetometers, Earth sensors, sun sensors and star sensors. However the interferometric sensor itself can be also a precise navigation sensor, i.e. attitude can be estimated by obtaining an RF signal generated by a ground station whose position is known. These two calibration methods are now under study.

4. Distributing Functions into Panels

In PETSAT, some satellite functions may be implemented as specialized functions of particular panels, and other functions may be implemented as standard functions in all types of panel. Therefore, one of the important research issues is how to distribute various satellite functions into different types of panel. To determine this, the following requirements for PETSAT features should be observed:

(a) Interface simplicity: panels of different types can be plugged-in in any quantity while satisfying the requirements on the four types of interface described in 2.2.

(b) Functions should be improved by increasing the number of panels: for example, the communication capacity should be reinforced

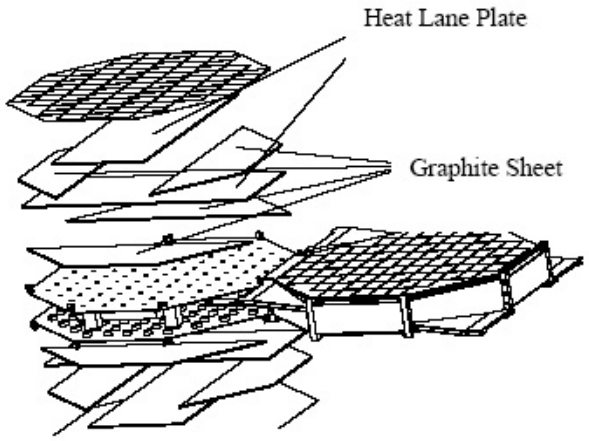


Figure 9: Thermal design inside the main panel structure

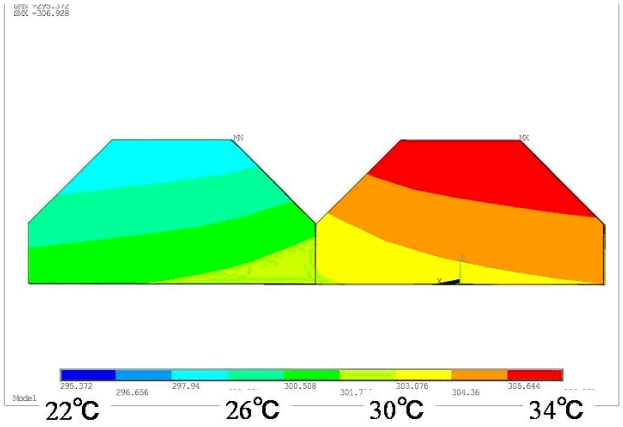


Figure 10: Results of thermal simulator for two connecting panels as modeled in Fig. 7

by the number of communication panels employed

(c) Standard panel structure: the structures of different panels should be almost the same so that mass production of the panel structure is possible.

(d) Flat panels: in order to be deployable, the panels should be flat, which requires that, for example, only one axis wheel can be implemented in one panel.

(e) Fault tolerance and graceful degradation: the satellite functions can be maintained at a degraded level in case of failures of certain panels or interface components (such as information lines, power lines, hinges, etc).

After the examinations taking into account how several example missions can be achieved by PETSAT, the following distribution strategy has been found appropriate.

(a) The variety of panels are as follows: ( ) show the specialized functions implemented in each panel

i. OBC Panel (high performance CPU and large memory)

ii. Communication Panel (transmitter, receiver and antenna)

iii. Attitude Control Panel (reaction/momentum wheel/magnetic torquer/gyro: one panel for one axis control)

iv. Orbit Control Panel (thruster, valve, pipes and propellant tank)

v. Battery Panel (large battery)

vi. Mission Panel (different for different missions)

(b) The following components are implemented as standard functions in all types of panel.

i. solar cell and small battery

ii. inter-panel communication-related components

iii. a local CPU and memory which manages information inside the panel

5. As to the electrical power, each panel's solar cell supplies power to the equipment in the panel, and the residual power can be transported to any panel through negotiation between the CPUs of each panel, considering the power demand of each panel

6. With respect to the information management, there are many CPUs, which should be efficiently utilized for realizing fault tolerance, graceful degradation, and grid computing to help in information processing for the mission panel

5. Example Designs of Panels and Interfaces

The PETSAT project is now in the initial conceptual study phase, during which the following first step design has been obtained.

5.1 Panel Structure

The panel structure should endure the launch environment. PETSAT is more tolerant against acceleration load because panels are stowed flat during launch. The vibration environment should, however, be considered seriously because the hinges between panels make the structural frequency lower. Another important requirement is that the same standard structure should be used for different types of panels. Considering the required mechanical interface described in 2.2, the main structure and the hinge-part, named link module should be separated as in Fig. 4. The link modules are of several types which can latch the panels at different angles (such as 90, 180, 270 °)

Fig. 5 shows one example design of the main structure, which consists of several plates and stiffener (placer between plates). The central plate, termed the mother board, has many holes aligned in a regular grid for attaching components to the panels. Fig. 6 shows one example of the interior design of a panel (for the OBC panel). The Bus Controller and battery are inserted as the standard functions, and CPU and large memory are the Specialized functions. Solar cells are not shown in this figure. In order to improve the heat conductivity within and between the plates, thermal conductors (carbon sheet and Heat Lane Plate) are inserted. Currently the size of the panel is designed as 300 mm by 300 mm with 50 mm height.

5.2 Thermal Design

According to the thermal interface requirements, the temperature difference between panels has been targeted to be less than 5°C. As the solar heat input is usually much larger than the power generated by internal components, the most severe situation, probably yielding the maximum temperature difference between panels, is the case where solar heat is incident on one panel (say, panel P), and the other panel (panel Q), which is 90 ° bent from panel P doesn't get any solar heat (Fig. 7). In this case, the 300 mm square panel P gets 135 W of solar heat, which should be transported to the other side of panel P and the neighboring panel Q.

Several heat conducting devices have been examined, and the two candidates, graphite sheet (heat conductivity is about 800W/m^{2}K) and small heat pipe named Heat Lane Plate (Fig. 8) are found to be appropriate for heat transfer within and between panels. The Heat Lane Plate is rigid structure and so cannot be used for inter-panel heat transfer. But, it has high heat transport capability (3×10^6 W/m^2). If only the graphite sheet is used for the inter-panel heat transfer, the intra-panel temperature difference is so large that it prevents efficient heat transfer between panels. So the Heat Lane Plate is inserted inside the panel to lower the intra-panel temperature difference. Fig. 9 shows the example design of inserting a Heat Lane Plate and graphite sheet inside the panel.

The thermal analysis using a simulator indicated that with these heat conducting devices, the intra- and inter-panel temperature difference can be less than 5°C. Fig. 10 shows the thermal simulator results for a simple model of two panels connected as in Fig. 7. It was shown that the temperature difference within one panel is less than 5 °C and the difference between the two panels at the connecting point can be suppressed under 5 °C.

5.3 Information Infrastructure

Firstly, the question of which of the two candidates for information lines, wired bus or RF LAN, should be employed has been studied, taking the interface requirements into account. Though RF LAN is superior in that no wiring is required, and so it is more tolerant to panel failure, the possible interaction between RF LAN and PETSAT-ground RF communication and the relatively large power required makes the usage of wired bus more attractive. With regard to the wired bus system, CAN bus is promising as it is capable of relatively high speed (1Mbps) and has already been used in many areas, so that there exists much supporting software/hardware for development of CAN bus system. Currently an information system using CAN bus and an SH-series bus controller is being developed.

6. Conclusions

A novel satellite concept of PETSAT is proposed, and its features, applications and basic design philosophy have been described. In order to further space development activities, satellite design philosophy needs an essential breakthrough so that reliable and capable satellites can be

fabricated within a shorter period and at far lower cost. PETSAT can be said to be one trial to approach such a breakthrough.

Acknowledgements

PETSAT research project is supported by NEDO subsidy.

Bibliography

[1] Yuya Nakamura, Hideyuki Tanaka, Masaki Nagai, and Ryu Funase. Multi-objective panel extension satellite petsat. In *10th Satellite Design Contest*, October 2002.