JBIS

Journal of the British Interplanetary Society

VOL. 66 No. 5/6 MAY/JUNE 2013

Contents

On the Possibility of Detecting Class A Stellar Engines using Exoplanet Transit Curves
Duncan H. Forgan

Asteroid Control and Resource Utilization
Graham Paterson, Gianmarco Radice and J-Pau Sanchez

Application of COTS Components for Martian Surface Exploration
Matthew Cross, Christopher Nicol, Ala' Qadi and Alex Ellery

In-Orbit Construction with a Helical Seam Pipe Mill
Neill Gilhooley

The Effect of Probe Dynamics on Galactic Exploration Timescales
Duncan H. Forgan, Semeli Papadogiannakis and Thomas Kitching

Innovative Approaches to Fuel-Air Mixing and Combustion in Airbreathing Hypersonic Engines
Christopher MacLeod

Gravitational Assist via Near-Sun Chaotic Trajectories of Binary Objects
Joseph L. Breeden

jbis.org.uk

ISSN 0007-084X

Publication Date: 24 July 2013

Notice to Contributors

JBIS welcomes the submission for publication of suitable technical articles, research contributions and reviews in space science and technology, astronautics and related fields.

Text should be:

1. As concise as the scientific or technical content allows. Typically 5000 to 6000 words but shorter (2-3 page) papers are also accepted called Technical Notes. Papers longer than this would only be accepted when the content demands it, such as a major subject review.
2. Mathematical equations may be either handwritten or typed, but must be clear. Unusual letters or symbols should be clearly identifiable. The solidus should be used for fractions in the text and for displayed formulae. Parentheses and fractional superscripts should be used to indicate square roots. On first use, acronyms should be translated in full into English. JBIS uses the SI system of units; if Imperial units are given, SI equivalent values should be added. The use of an Equation Editor is strongly advised.
3. References and notes (i.e. replacement for end notes) should be shown in the text in numerical sequence in square parentheses e.g. [3], with the full list at the end of the paper in the following styles:

   1. A. June and W.R. Camp, "Space Trajectories", *J. Aerospace Sci.*, **20**, pp.635-642, 1983. (Journal Paper – where 20 is the volume number and pp gives page numbers)
   2. N. Sisakyan and V. Yazdavdkiy (eds), "First Group Flight into Outer Space", Nauka Publishing House, Academy of Sciences, Moscow, 1964. (Book)
   3. D.F. Lawden, "Interplanetary Rocket Trajectories", in *Advance in Space Science*, **Vol. 1**, ed. F.I. Ordway III, Academic Press, New York, p.3, 1959. (Article in Book)
   4. D.M. Ashford, "The Aeroplane Approach to Launch Vehicle Design", *JBIS*, **37**, pp.163-71, 1984. (Paper in JBIS)
4. Internet sources are not generally acceptable as scientific confirmatory evidence, and are often transitory in nature. Their use as reference material should be carefully evaluated in the context of the contention they are being used to support. If they are used include the date the site was accessed in the following style: (Last Accessed 12th April 2012).

Illustrations:

1. Each figure should be cited in numerical order in the text. Figure captions should be short and concise. Every figure used must be referred to in the text.
2. Permission should be obtained and formal acknowledgement made for use of a copyright illustration or material from elsewhere.
3. **Photos:** Original photos are preferred if possible or good quality copies. Photos should be clear, with good contrast and free from blemish. If this is not possible please scan them with a resolution of 350 dpi.
4. **Line drawings:** Original drawings are preferred. If this is not possible please scan them with a resolution of 600 dpi. Photocopies are not usually of an acceptable quality.
5. **Illustration size:** please ensure that all electronically submitted illustrations are a minimum width of 140 mm.
6. Images embedded in electronic documents are acceptable on first submission for review purposes, however if the paper is accepted for publication each illustration will be required as a separate file.

Paper Submission:

1. JBIS requires papers to be submitted in an electronic form compatible with PC formats. We can handle Microsoft Word for Windows up to version Word 2010. A PDF version of the paper would also be helpful to aid formatting.
2. For preference we would ask author's to submit manuscripts via email to: jbis@bis-spaceflight.com. However if this is not possible, manuscripts can also be submitted by post and sent to the JBIS office address below.
3. A License to Publish form is required. This can be sent separately from the paper but is required before the paper will be published. It is the corresponding author's responsibility to ensure all other authors are in agreement with the submission and the terms of publication. A License to Publish can be downloaded from - www.bis-space.com/wp-content/uploads/2012/08/WebsiteLicense.pdf.

Authors will receive a complimentary copy of the issue in which their paper appears.

Our full Guidelines for Authors can be downloaded from - www.bis-space.com or email JBIS@bis-spaceflight.com.

PRODUCTION ASSISTANT:

B. Jones

PROMOTION:

S. Parry

JBIS Office:

27/29 South Lambeth Road, London,
SW8 1SZ, England.

Tel: 020 7735 3160
Fax: 020 7587 5118
Email: jbis@bis-spaceflight.com
www.bis-space.com

DISTRIBUTION DETAILS

JBIS is distributed worldwide by mail and may be received by annual subscription or purchase of single copies. It is available through membership of the British Interplanetary Society at much reduced rates. Subscription details for members, non-members and libraries are available from the above address.

JBIS is a publication which promotes the objectives of the British Interplanetary Society. Opinions expressed in signed articles are those of the contributors and do not necessarily reflect the views of the Editor or the Council of the British Interplanetary Society. Security clearance, where necessary, is the responsibility of the author.

* * *

Published by the British Interplanetary Society. Registered Company No: 402498. Registered Charity No: 250556.

Printed by Latimer Trend & Company Ltd, Estover Road, Plymouth, PL6 7PY, England.

* * *

JBIS

Journal of the British Interplanetary Society

VOL. 66 No. 5/6 MAY/JUNE 2013

LIST OF CONTENTS

143 **Editorial**

144 **On the Possiblity of Detecting Class A Stellar Engines using Exoplant Transit Curves**
Duncan H. Forgan

155 **Asteroid Control and Resource Utilization**
Graham Paterson, Gianmarco Radice and J-Pau Sanchez

161 **Application of COTS Components for Martian Surface Exploration**
Matthew Cross, Christopher Nicol, Ala' Qadi and Alex Ellery

167 **In-Orbit Construction with a Helical Seam Pipe Mill**
Neill Gilhooley

171 **The Effect of Probe Dynamics on Galactic Exploration Timescales**
Duncan H. Forgan, Semeli Papadogiannakis and Thomas Kitching

178 **Innovative Approaches to Fuel-Air Mixing and Combustion in Airbreathing Hypersonic Engines**
Christopher MacLeod

190 **Gravitational Assist via Near-Sun Chaotic Trajectories of Binary Objects**
Joseph L. Breeden

* * *

EDITORIAL

Dear Readers,

A warm welcome to you all as the summer approaches and you look forward to your issues of *JBIS* dropping through the letter box. I have mentioned this in the past but I want to do so again, to remind you that *JBIS* does receive letters from readers pertaining to past published papers. I feel this is important as it continues the peer review process and ensures effective scrutiny of people's research, whilst also opening up promising new lines of thought. So, please feel free to drop us a technical letter. Whilst on the subject of reminders, please do also have a look at the *JBIS* web site (www.jbis.org.uk). It has come a long way since it first launched and you can now download papers for a small fee. Currently, we are also working on an online editorial system. We are also looking at the possibility of receiving LaTeX submissions. None of this will happen overnight, but we are gradually making progress to improve the oldest astronautical journal in the world.

Now for this issue which contains a mix of papers, as an interlude to the 100YSS papers we have been publishing recently. First is a paper by Duncan Forgan on the possibility of detecting class A stellar engines using exoplanet transit curves. Next we have a paper by Graham Paterson and colleagues which explores the potential for asteroid control and resource utilization. The authors argue that asteroids are materials rich small solar system bodies which are candidates for rendezvous and mining. The authors examine the scenario of asteroid capturing in detail. Matthew Cross and associates write about the application of commercially available off the shelf components for Martian surface exploration.

Neil Gilhooney has written an intriguing paper discussing in-orbit construction with helical seam pipe mill. This approach is meant to be a novel way to address the construction of large structures in space and in particular a torus habitat. In another paper from Duncan Forgan and colleagues, he explores the effect of probe dynamics on galactic exploration timescales. The authors present multiple realisations of single probes exploring a small patch of the Milky Way. Chris MacLeod examines innovative approaches to fuel-air mixing and combustion in airbreathing hypersonic engines. Finally, we have a thought provoking submission from Joseph Breedon, exploring the possibility of interstellar travel with Binary Asteroids. By selecting the correct chaotic trajectory the author calculates the velocity that may be achieved if this mission can be made plausible.

We have several other special issues in the planning, including a paper dedicated to the research of Dr Les Shepherd, papers from the 2013 Tennessee Valley Interstellar Workshop. We also have papers pending from a recent Chinese/Soviet symposium and a symposium discussing the philosophy of the starship. Recently, we in the *JBIS* team put together a feedback form so we can see how we are doing. You can find this form in this issue (middle pages) or online at www.surveymonkey.com/s/Q5XTT7M. Please complete it if you get time and let us know how we are doing. Please enjoy the latest issue of your journal.

Kelvin F. Long, Editor *JBIS*

JBIS, Vol. 66, pp.144-154, 2013

ON THE POSSIBILITY OF DETECTING CLASS A STELLAR ENGINES USING EXOPLANET TRANSIT CURVES

DUNCAN H. FORGAN
Scottish Universities Physics Alliance (SUPA), Institute for Astronomy, University of Edinburgh, Blackford Hill, Edinburgh, EH9 3HJ, UK.
Email: dhf@roe.ac.uk

The Class A stellar engine (also known as a Shkadov thruster) is a spherical arc mirror, designed to use the impulse from a star's radiation pressure to generate a thrust force, perturbing the star's motion. If this mirror obstructs part of the stellar disc during the transit of an exoplanet, then this may be detected by studying the shape of the transit light curve, presenting another potential means by which the action of extraterrestrial intelligence (ETI) can be discerned. We model the light curves produced by exoplanets transiting a star which possesses a Shkadov thruster, and show how the parameters of the planet and the properties of the thruster can be disentangled provided that radial velocity follow-up measurements are possible, and that other obscuring phenomena typical to exoplanet transit curves (such as the presence of starspots or intrinsic stellar noise) do not dominate. These difficulties aside, we estimate the *a priori* probability of detecting a Shkadov thruster during an exoplanet transit, which even given optimistic assumptions remains stubbornly low. Despite this, many exoplanet transit surveys designed for radial velocity follow-up are on the horizon, so we argue that this remains a useful serendipitous SETI technique. At worst, this technique will place an upper limit on the number of Class A stellar engines in the Solar neighbourhood; at best, this could help identify unusual transiting exoplanet systems as candidates for further investigation with other SETI methods.

Keywords: SETI, exoplanet, transits, megastructures

1. INTRODUCTION

For much of its sixty-year history, the Search for Extraterrestrial Intelligence (SETI) has relied heavily upon the detection of artificial radio signals emitted in the frequency range often referred to as the "Water Hole" or terrestrial microwave window [1], a band of frequencies where the Earth's atmosphere is transparent and interference from Galactic and cosmic background radiation is minimal. This band also happens to contain the 21 cm (or 1.42 GHz) spectral line emitted by neutral hydrogen atoms when the alignment of proton-electron spins "flip" from parallel to anti-parallel. As hydrogen is the most abundant element in the Universe, it is often argued that extraterrestrial intelligences (ETIs) with radio technology will be scanning frequencies near to the 21 cm line, and may even choose to emit signals at frequencies related to it (e.g. 21π cm or $21/\pi$ cm). This would suggest that Earth-like intelligent species could be detected by searching frequencies in the Water Hole for narrowband transmissions or broadband pulses [2].

This motivation has inspired many radio SETI surveys carried out to date, placing constraints on the number of civilisations emitting in the surveyed bands in the Solar neighbourhood. Most recently, a SETI survey of 86 stars in the Kepler field known to host transiting exoplanets estimates that less than 1% of the targets host civilisations that emit in the 1-2 GHz band, at strengths comparable to those producible by humankind [3]. Recent SETI searches using Very Long Baseline Interferometry (VLBI [4]) show that the next generation of telescope arrays such as the Square Kilometre Array (SKA) will be able to survey the local neighbourhood to even higher fidelity, while efficiently discriminating potential signals from human radio frequency interference (see also [5]).

Equally, SETI scientists have recognised that searching in a limited region of the electromagnetic spectrum may be counter-productive to the overall goal of detection. Even at the beginning of radio SETI, some argued for searches in the optical [6], noting that technology akin to the then recently-invented laser would be a highly efficient means of interstellar communication, both in its potential to traverse much larger distances at the same energy cost, and the ability to encode more information per unit time due to their higher carrier frequency. While extinction by interstellar dust could destroy or mask a signal, ETIs can strike a balance by reducing the carrier frequency and emitting in the near infrared [7]. Recent searches for laser pulses in the spectra of local main sequence stars (e.g. [8, 9]) are also yet to bear fruitful detections.

While both radio and optical SETI have their orthogonal advantages and disadvantages, they share a weakness common to all transient astrophysics - repeatability of detection. The few cases where SETI searches uncovered a potential signal (e.g. the "Wow!" signal of 1977 [10, 11]), the candidate signal was not rediscovered. In the most memorable case where an apparently artificial signal was detected repeatedly, it was soon shown to be a previously undiscovered natural phenomenon - the pulsar [12, 13]. As pulsing transmitters are generally more energy efficient, there is a strong element of serendipity in detecting intentional transmissions, and as such surveys must be carefully designed [14, 15]. While it is possible that ETIs may deliberately repeat their signals to make their acquisition easier, they are under no compulsion to do so.

If the goal is to detect unintentional transmissions, then this obstacle can be avoided, as these transmissions are expected to

be produced as part of the typical activities of an extraterrestrial civilisation. Unfortunately, unintentional transmissions by their very nature are not designed to travel large interstellar distances - terrestrial radio leakage could be detected by current and future instrumentation at no more than a few hundred parsecs [16], and their content is likely to be less informative than that transmitted by a dedicated beacon. Indeed, in the case of radio signal leakage, it may only be possible to detect civilisations in their earliest uses of radio communication, as they may wish to subsequently reduce signal leakage as a cost-cutting measure [17], or to shield themselves from dangerous eavesdroppers [18, 19].

However, just as a planetary biosphere displays atmospheric biomarkers that can be used in some cases to infer the presence of life [20, 21], planetary systems inhabited by intelligent species are likely to present *noomarkers* (or sometimes "technomarkers") which are difficult to hide. Using the classification system for civilisations proposed by [22], these noomarkers are produced strongly by Type II civilisations, which manipulate energy resources equivalent to typical stellar luminosities, rather than the Type I civilisations more reminiscent of humankind, whose access to energy resources is limited to the much smaller planetary scale. Having said this, we should note that noomarkers produced by asteroid mining [23], may be present in the Solar System, albeit weakly, before human civilisation reaches the Type II classification, if recently founded commercial ventures to mine asteroids are successful (see e.g. http://deepspaceindustries.com and www.planetaryresources.com).

Megastructures on the scale of celestial objects constitute the most studied type of noomarker, the classic example being the Dyson sphere [24]. Indeed, SETI searches for this type of artifact [25, 26, 27] are often referred to as "Dysonian SETI" in deference to its most well-known advocate [28]. Megastructures or macro-engineering projects are motivated by a Malthusian view of the future development of humanity. Rising demand from an increasing population eventually encounters the paucity of supply, as terrestrial resources are exhausted. When this happens, the high costs of macro-engineering projects can soon be outweighed by the enormous benefits they can bring, not to mention the predicted costs of *not* carrying out the project [29]. This is hardly an ironclad proof that ETIs will always build megastructures, and one can be easily led astray when speculating on the motives of advanced technological civilisations, but at the very least it suggests that there are some evolutionary trajectories for technological civilisations that include the construction of Dysonian-type artifacts, especially if a civilisation adopts an optimisation-driven rather than expansion-driven developmental strategy [30].

There have been many suggestions of other noomarkers that are amenable to SETI searches, such as pollution in planetary [31] and stellar atmospheres [32], or deliberate "salting" of the host star with unusual isotopes (e.g. [33, 34]). These suggestions are motivated by the fact that there is a great deal of study of stars and planets for their own sake.

From the first detection of an extrasolar planet (or exoplanet) around a main sequence star almost two decades ago [35], there has been a drive to discover and characterise a large number of exoplanets, as the properties of the exoplanet population allows astronomers to understand the physical processes at work in planet formation. While exoplanets can be detected in a variety of ways (see [36] for a general review), the majority of the known exoplanet population has been detected using one or (or both) of the following methods. The *radial velocity* method (cf [37]) studies the star's spectral lines for evidence of changes in velocity along the line of sight to the observer. As stars and planets orbit a common centre of mass, the evolution of the star's radial velocity encodes information regarding the planet's orbital period and semimajor axis, as well as the minimum mass of the planet, $M_p \sin i$, where M_p is the true mass and i is the inclination of the orbit (with $i = 90^o$ equivalent to the star, planet and observer existing in the same plane). The transit method (cf [38]) requires the observer to study the star's flux as a function of time. If the planet passes between the star and the observer, this produces a dimming of the stellar flux proportional to the size of the silhouette cast by the planet on the stellar disc. The transit method constrains the planet's physical radius r_p relative to the stellar radius R_*. The combination of both methods allows the inclination to be measured, and the true mass M_p to be determined.

While other detection methods exist, such as gravitational microlensing (e.g. [39]) or direct imaging (e.g. [36]), radial velocity and transit studies have not only revealed the largest quantity of exoplanets to date, but also the most detailed data on individual exoplanet systems, when both methods are used in tandem.

Exoplanet detection benefits SETI in several ways; firstly, it provides invaluable data on the frequency of potentially habitable terrestrial planets, constraining one of the seven (or more) terms of the master equation of SETI, the Drake Equation, which is cast in various guises [40, 41, 42, 43]. Secondly, exoplanet studies can in principle identify objects which are highly likely to not be merely habitable, but inhabited. For example, exoplanet transit spectroscopy, which uses multi-wavelength observations of transits to study the planet's atmospheric absorption as a function of wavelength, is likely to prove effective as a probe for the presence of biomarkers ([44] and references within). Thirdly, it provides an extensive, rich dataset of astrophysical phenomena on planetary system scales, which can be studied in a serendipitous or "piggyback" fashion for evidence of artificial signals or noomarkers.

In particular, the transit detection method allows the exciting possibility of indirectly detecting megastructures through their obscuration of the parent star. [45] illustrates how transit curves could contain a deliberate signal placed by ETIs in the form of large geometric structures, which have transit signatures distinct to that of an occulting spherical disc. Such structures could be placed in an orbit where observing civilisations would be able to easily detect an exoplanet transit (cf [46]), with the lifetime of the resulting signal potentially very long indeed [47].

In this paper, we focus on a particular class of megastructure, known as the Class A stellar engine, or the Shkadov thruster, in honour of its first description by [48]. This object is essentially a mirror which reflects a fraction of the star's radiation pressure, causing a force asymmetry which exerts a thrust on the star. This thrust could be used to move a civilisation's host star from its "natural" orbit if it posed some harm to the civilisation (e.g. a dangerous close approach to another star or dust cloud). We study the potential transit signatures produced by an exoplanet in a system that contains such an engine, and show how to disentangle the planet's contribution to the transit from the engine's occultation of the star. In section 2 we describe the physics of the Shkadov Thruster; in section 3 we

recap the structure of a transit curve in the absence of a stellar engine; in section 4 we show how the transit curve is modified in the presence of a Shkadov thruster, and how to analyse these curves in section 5; in section 6 we discuss the implications of this work and we summarise our conclusions in section 7.

2. STELLAR ENGINES AND THE SHKADOV THRUSTER

In short, a stellar engine is any device or structure which extracts significant resources from a star in order to generate work. Typically, the resource in question is the radiation field produced by the star, although some stellar engines use the mass of the star as propellant, e.g. the stellar rockets or "star lifters" as proposed by [49] and [50].

Strictly speaking, the most well known theoretical megastructure, the Dyson sphere [24], is not a stellar engine, as it does not specifically generate work. The Dyson sphere is a spherical shell that englobes the star, with a radius and thickness such that the radiation pressure force from the star and the gravitational force on the sphere are in balance, producing a static satellite (or *statite*). While usually envisaged as a single solid structure, it can be considered as a very large collection of smaller bodies, arranged in a spherical geometry. This is sometimes referred to instead as a Dyson swarm, and is used only to collect energy.

The Dyson sphere/swarm's function is therefore to either a) collect stellar energy, or b) provide living space, or both. The interior of a single solid Dyson sphere, designed as an Earth-type habitat, will provide a significantly larger amount of living space compared to a planetary body, as the radius of a Dyson Sphere will be very large relative to terrestrial planet radii.

While the Dyson sphere is not a stellar engine, most stellar engine designs share similarities with the Dysonian model - in particular, they share the Dyson sphere's hunger for construction material. A full Dyson sphere would require the systematic destruction of several planetary bodies to provide raw materials [24, 51]. Badescu and Cathcart [51] present a classification system for stellar engines: we will focus exclusively on their Class A stellar engine, or Shkadov thruster (see Fig. 1).

A spherical arc mirror (of semi-angle ψ) is placed such that the radiation pressure force generated by the stellar radiation field on its surface is matched by the gravitational force of the star on the mirror.

Radiation impinging on the mirror is reflected back towards the star, preventing it from escaping. This force imbalance produces a thrust of magnitude [51]

$$F = \frac{L}{2c}(1-\cos\psi) \tag{1}$$

where L is the stellar luminosity. This result is similar to that produced by [48], although their calculation assumes the temperature of the star remains constant. In reality, the reflected radiation will alter the thermal equilibrium of the star, raising its temperature and producing the above dependence on semi-angle. Increasing ψ increases the thrust, as expected, with the maximum thrust being generated at $\psi = \pi$ radians. However, if the thruster is part of a multi-component megastructure that includes concentric Dyson spheres forming a thermal engine, having a large ψ can result in the concentric spheres possessing poorer thermal efficiency [51].

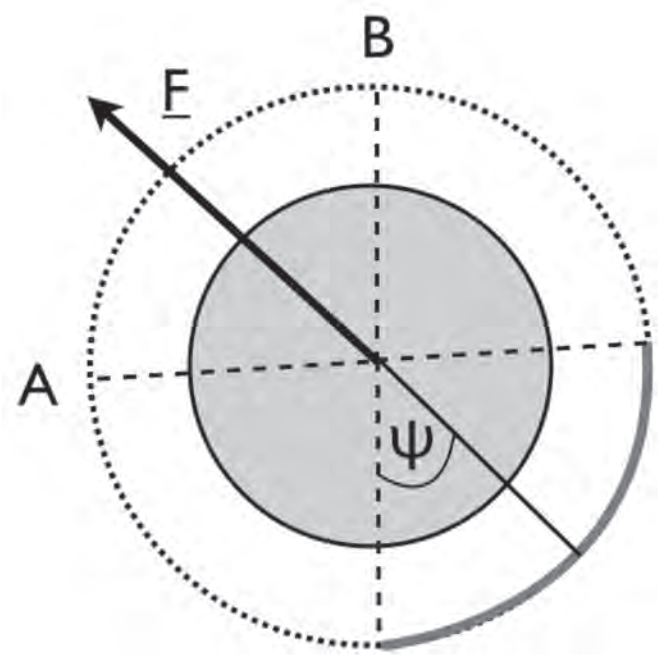


Fig. 1 Diagram of a Class A Stellar Engine, or Shkadov thruster. The star is viewed from the pole – the thruster is a spherical arc mirror (solid line), spanning a sector of total angular extent 2ψ. This produces an imbalance in the radiation pressure force produced by the star, resulting in a net thrust in the direction of the arrow.

3. EXOPLANET TRANSIT CURVES

We briefly recap the typical shape of an exoplanet transit curve here (for a review see e.g. [38]) Rather than attempting to simulate observations by pixelating the stellar disc and carrying out parameter fitting, as was done by [45], we instead construct theoretical curves in the following fashion.

We work in (x, y) coordinates, where the coordinates represent the projection of the planetary system onto a two-dimensional "observer" plane. A single star is placed at the origin (0, 0), possessing a stellar disc with total area $A_* = 1$, i.e its radius is

$$R_* = \sqrt{\frac{A_*}{\pi}} = \sqrt{\frac{1}{\pi}} \tag{2}$$

in arbitrary units. A planet (with a disc area $A_p = 0.01$, and radius $r_p = \sqrt{0.01/\pi}$) is positioned initially at (x_i, b), where b is the impact parameter. We assume that the projection of the planet's orbit on the plane is a straight line rather than an arc, and move the planet in the x-axis only, across the stellar disc until it reaches (x_f, b). As transit curves typically normalise the received flux by the steady value obtained during a non-transiting period of the planetary orbit, we calculate the flux using

$$F = 1.0 - \frac{A_{int}}{A_*} \tag{3}$$

Where A_{int} is the area inside the intersection of the stellar and planetary discs, and $A_{int} = [0, A_p]$. By construction, the minimum

flux is 0.99, and we neglect thermal emission from the planet. A_{int} is analytically calculable provided the positions of the star and planet are known, by solving a quartic equation for the intersection points of the two discs, and calculating the circular segment area subtended inside each disc between the intersection points.

The transit curve can be deconstructed into several components, delineated by four events where the planetary and stellar discs possess one intersection point only, i.e. the discs are in tangential contact:

1. Beginning of Transit Ingress (τ_I) - where the star is as yet unaffected by transit.
2. End of Transit Ingress (τ_{II}) - the planet is now fully within the stellar disc. In the absence of stellar limb-darkening, the light curve would reach its minimum value here.
3. Beginning of Transit Egress (τ_{III}) - the planet is nowmoving outside the stellar disc, and the received stellar flux increases, and finally
4. End of Transit Egress (τ_{IV}) - the planet no longer obscures the stellar disc, and the transit is complete.

Figure 2 shows the transit curves produced both in the absence and presence of stellar limb-darkening. We do not model the curves as a function of time, but rather as a function of planetary position. We use a dimensionless phase parameter ϕ for the x-axis in this and all subsequent curves, where $\phi = x_p/R_*$, and x_p is the x-coordinate of the planet's centre. In these units, the centre of the planet touches the stellar perimeter at $\phi = -1, 1$.

To simulate limb-darkening, we modify A_{int} using

$$A_{int} \rightarrow A_{int}\left(1-u\left(1-\mu\right)\right) \tag{4}$$

where

$$\mu = \sqrt{\frac{R_*^2 - d^2}{R_*^2}} \tag{5}$$

the distance between the centres of the stellar and planetary discs is denoted by d, and we select the limb-darkening parameter $u = 0.6$ (where u is typically a function of the observing wavelength). Equation 4 is one of many different expressions used for limb-darkening which are derived from stellar-atmosphere models (see [52] for a list), but this work is not particularly sensitive to which parametrisation that we adopt, but merely requires that we adopt one. Increasing the dimensionless impact parameter

$$b = \frac{a\cos i}{R_*} \tag{6}$$

can partially mimic the curve shape in the ingress-egress sections that is also produced by the limb-darkening. However, both parameters can usually be disentangled with an appropriate best-fit modelling procedure (see e.g. [38]). If the oblateness of the planet is non-zero, or the planet possesses rings, then this can be detected provided the photometric precision is better than ~10^{-4} [53, 54].

Let us define the transit depth (i.e. the fractional difference in flux before and during the transit) as d and the ratio of the planet to stellar radius $r_p/R_* = k$. As such, if both the stellar and planetary discs are circular,

$$\delta = k^2 \tag{7}$$

While in this work we model the transit curve as a function of planetary position, not time, it will be useful to define the equatorial crossing timescale (for a circular orbit):

$$\tau_{eq} = \frac{R_* P_{pl}}{2\pi a_p} \tag{8}$$

Where P_{pl} is the planet's orbital period, and a_p is the planet's semi-major axis. The total transit time is

$$T_{tot} = \frac{P}{\pi}\sin^{-1}\left(\frac{R_*}{a}\sqrt{1+k^2-b^2}\right) \tag{9}$$

We can therefore convert our dimensionless phase into real time by using

$$t = \frac{\phi T_{tot}}{2} \tag{10}$$

where $t = 0$ represents the midpoint of the transit. If we assume

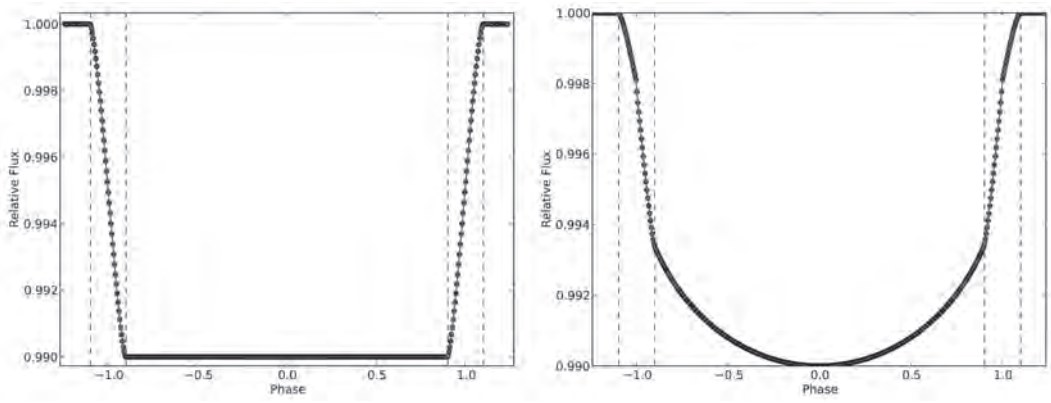


Fig. 2 Examples of the transit curves produced for a planet without a stellar engine present (with impact parameter b = 0). The left panel shows the curve in the absence of stellar limb-darkening, and the right panel shows the curve with limb-darkening. The dotted lines in each plot represent (from left to right): beginning of ingress (τ_I), end of ingress (τ_{II}), beginning of egress (τ_{III}) and end of egress (τ_{IV}).

that $r_p << R_* << a$, and $b << 1 - k$ (i.e. we exclude transits that come close to grazing the upper or lower limb of the star), the ingress timescale is:

$$\tau_{ing} \equiv \tau_{II} - \tau_I = \tau_{eq}\left(\sqrt{(1+k)^2 - b^2} + \sqrt{(1-k)^2 - b^2}\right) \sim 2\tau_{eq}\frac{k}{\sqrt{1-b^2}} \quad (11)$$

Armed with an appropriate model to fit stellar limb darkening and remove sources of false positives such as starspots, the transit curve, specifically the parameters (δ, T_{tot}, τ_{ing}), is sufficient to characterise k and b. Stellar atmosphere modelling, or in some cases asteroseismology [55] can then be used to define R_* to obtain $r_p = kR_*$.

4. TRANSIT CURVES IN THE PRESENCE OF A CLASS A STELLAR ENGINE

We model the presence of an obscuring Class A Stellar Engine as a circular segment drawn on the stellar disc at an angle ξ to the y-axis, passing through some point on the stellar disc (β_1, β_2) (see Fig. 3). In practice, we will measure these values normalised to the stellar radius in the same manner as the transit's impact parameter, i.e. β_1, β_2 = [0, 1].

The engine completely obscures the star to the right hand side of the chord, and as a result when the planetary disc passes into the region where the star is obscured, the transit curve will be altered, particularly its egress features. Note that throughout this work we position the engine in the positive x-axis region of the star without affecting the generality of the result - engines in the negative x region would simply produce signatures at the curve's ingress point rather than the egress point.

We define the obscured area of the stellar disc as Σ_s. As the engine is a highly reflective mirror, there is no limb-darkening associated with the obscured side of the star. The depth of the transit curve will no longer be equivalent to the ratio of the planetary and stellar disc areas, but will also depend on Σ_s:

$$F = 1.0 - \frac{A_{int} - A_{int,s}}{A_* - \Sigma_s} \quad (12)$$

where $A_{int,s}$ is the area inside the intersection of the planetary disc and the Shkadov thruster. When the planet passes into the region where the stellar disc is obscured, the transit curve will show the planet beginning "egress" earlier than in the case where the engine is absent, introducing an asymmetry (see Fig. 4). We will hereafter refer to transits of this type as "Shkadov transits", and we will refer to the apparently early egress of the curve as the "Shkadov egress". Although it is strictly also correct to refer to it as the Shkadov ingress, it is perhaps more sensible to allow each curve to have an ingress and egress component rather than two of the same. As the ingress section of the curve is limb-darkened and the Shkadov egress is not, the ingress and egress will progress with different measured gradients.

While these curves of this type are difficult to produce naturally, and as such are *prima facie* evidence for macro-engineering occuring in this particular star system, what else can be divined from the curve? Specifically, can we determine

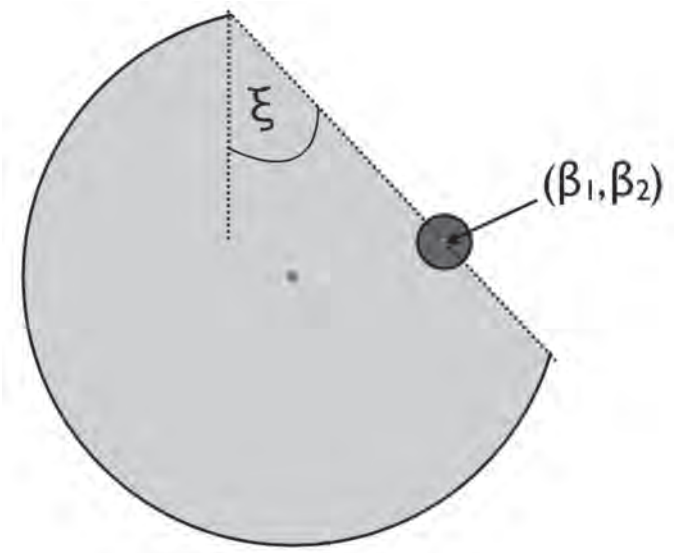


Fig. 3 Diagram of a transit in the presence of a Class A Stellar Engine or Shkadov Thruster (which we will refer to throughout as a "Shkadov transit"). The Shkadov thruster subtends a chord at angle x to the y-axis, and obscures an area of the stellar disc Σ_s, in this case to the right of the chord. The chord passes through a point (β_1, β_2) on the stellar disc (in the diagram, $\beta_2 = b$, the planet's impact parameter). As the thruster is a highly reflective mirror, there is no limb darkening associated with it.

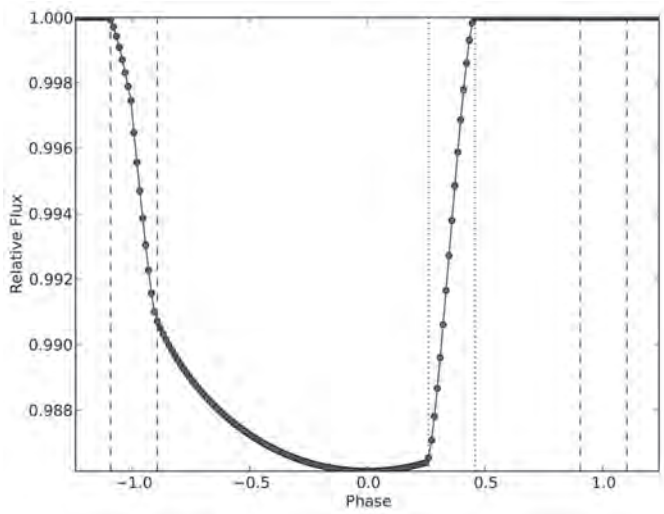


Fig. 4 The light curve associated with a Shkadov Transit. The dashed lines represent the expected ingress/egress points for a transit without a Shkadov thruster present, and the dotted lines represent the beginning and end points of the planet's ingress into the region obscured by the Shkadov thruster, which we refer to as the "Shkadov egress".

the properties of both the stellar engine as well as the true exoplanet parameters? After all, the curve's symmetry is lost. We cannot measure T_{tot}; the depth of the Shkadov transit curve, $\tilde{\delta}$, is now degenerate in r_p and Σ_s:

$$\tilde{\delta} = \frac{k^2}{1 - \Sigma_s \Big/ \left(\pi R_*^2\right)} \quad (13)$$

In other words, what can we now measure to recover (k, b) and the properties of the thruster?

We can see the beginning of the solution to this problem by

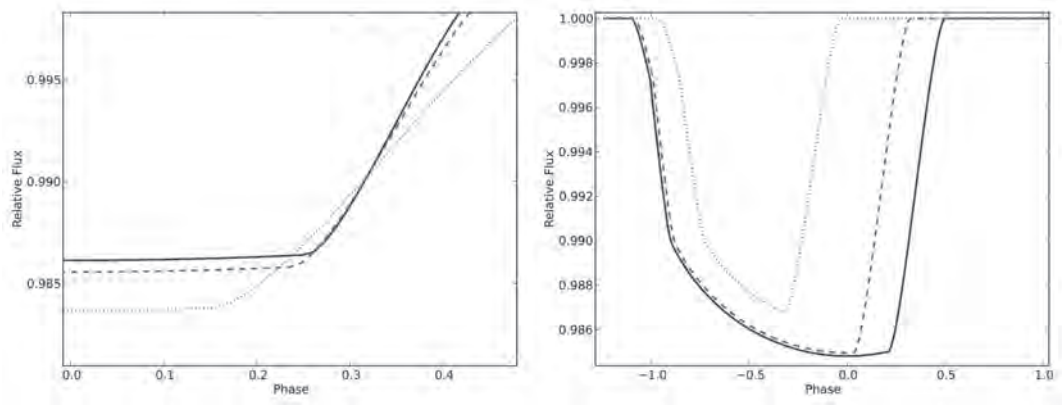


Fig. 5 (Left) The egress section of the light curve during Shkadov transits for a planet with impact parameter b = 0, as the angle between the engine and the y-axis, ξ, is increased. The dotted curve (with the deepest transit depth) has $\xi = 0^o$; the dashed curve with moderate depth has $\xi = 30^o$, and the shallow solid curve is $\xi = 60^o$. (Right) The dependence of the Shkadov transit light curve as the planetary impact parameter, b, is increased, for an engine of fixed $\xi = 45^o$ and $(\beta_1, \beta_2) = (0.2, 0)$. The solid curve (with the longest transit duration) has b = 0; the dashed curve with moderate duration has b = 0.1, and the dotted curve with shortest duration has b = 0.3.

holding (β_1, β_2) fixed and varying ξ (left hand panel of Fig. 5). The midpoint of ingress into the Shkadov thruster is fixed at β_1 (note all curves intersect at a phase equal to β_1). Increasing ξ while holding (β_1, β_2) fixed reduces Σ_s, and the duration of the Shkadov egress also increases.

If we now consider the engine having fixed values of $\xi = 45^o$ and (β_1, β_2), and instead vary the planet's impact parameter b, then we can see (right hand panel of Fig. 5) how the presence of the engine will cut the transit short depending on its value of b. Planets with higher b will encounter the obscured area earlier, and hence egress begins earlier. As we model the silhouette of the engine's edge as a straight line, all the curves in the right hand panel of Fig. 5 show the same duration of egress.

Provided that:

i. ingress of the exoplanet into the stellar disc is not also obscured,

ii. radial velocity follow-up measurements are possible, and

iii. the radius of the star R_* is still measurable (through e.g. asteroseismology measurements),

this should be sufficient to estimate the transiting timescale τ_{eg}, the ingress timescale τ_{ing} and hence the degenerate $k/\sqrt{1-b^2}$.

If the orbital period P_{pl} and the semi-major axis a_p are well-constrained as a result of the combined radial velocity/transit measurements, then the transit curve should be sufficiently informative to break the degeneracies linking (r_p, ξ, β_1, β_2). The measurements to be made are: the Shkadov egress timescale $\tilde{\tau}_{eg}$, the Shkadov transit depth $\tilde{\delta}$, and the midpoint of Shkadov egress $\phi_{eg} = \beta_1$. At the midpoint of Shkadov egress, $\beta_2 = b$, reducing the number of parameters to estimate by one.

5.1 The Shkadov Egress Timescale

As can be seen from Fig. 5, the time taken for the planet to complete Shkadov egress is independent of b, and depends only on ξ. If we define the Shkadov thruster as a circular segment of central angle θ, and height h (see Fig. 6), then we can analytically calculate how the area of the planetary disc obscured by the thruster changes as the planet sweeps across the chord.

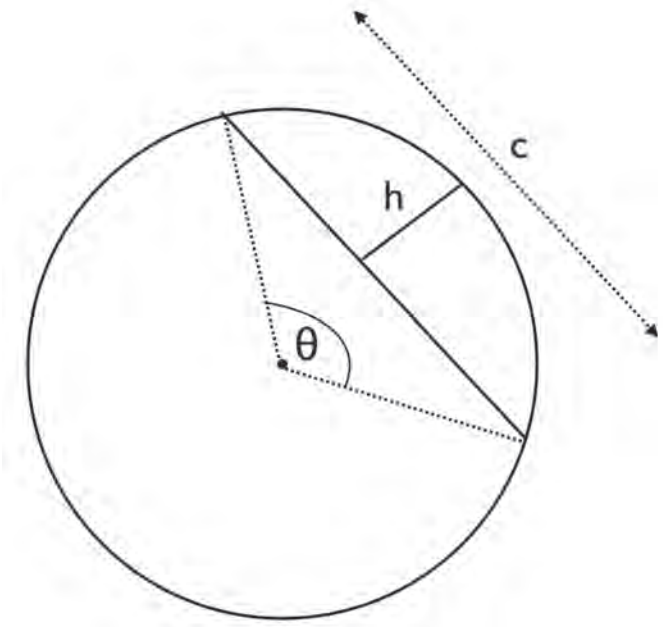


Fig. 6 Diagram of a circular segment, with central angle θ, height h and chord length c.

The height h of a circular segment (inside a circle of radius R) and the central angle that spans the segment, θ are related by:

$$h = R\left(1-\cos\left(\frac{\theta}{2}\right)\right) \tag{14}$$

The length of the chord defining the segment is

$$c = 2R\sin\left(\frac{\theta}{2}\right) \tag{15}$$

and the area of the segment is

$$A(\theta) = \frac{R^2}{2}(\theta - \sin\theta) \tag{16}$$

As the planetary disc sweeps across the thruster chord, the thruster obscures a circular segment of height $h_p(t)$ and angle $\theta_p(t)$. If the thruster has $\xi = 0^o$, then h_p is parallel to the x axis and

$$\frac{dh_p}{dt} = \dot{h}_p = \dot{x} \tag{17}$$

It can be shown that if $\xi \neq 0^o$, then

$$\dot{h}_p = \dot{x}\cos\xi \tag{18}$$

Differentiating equation (14) gives

$$\dot{h}_p = \left(\frac{R_*}{2}\sin\theta_p\right)\dot{\theta}_p = \frac{c}{4}\dot{\theta}_p \tag{19}$$

Applying equation (16) to the planetary disc, and taking the first derivative gives

$$\dot{A} = \frac{r_p^2}{2}\left(1-\cos\theta_p\right)\dot{\theta}_p \tag{20}$$

To solve this, we must solve for $\dot{\theta}_p$. From equation (19):

$$\dot{\theta}_p - \frac{4\dot{h}_p}{c} = \dot{\theta}_p - \frac{4\dot{h}_p}{2r_p\sin\left(\frac{\dot{\theta}_p}{2}\right)} = 0 \tag{21}$$

Differential equations of the form

$$\dot{y} - \frac{A}{\sin\left(\frac{y}{2}\right)} = 0 \tag{22}$$

Have the solution

$$y(t) = \pm 2\cos^{-1}\left(\frac{1}{2}\left(-At - C\right)\right) \tag{23}$$

where C is a constant of integration. If we impose the initial condition $y(0) = 0$, then $C = -2$. Substituting for $A = 2\dot{h}_p/r_p$, and assuming $\dot{x}$ is constant:

$$\cos\frac{\theta_p(t)}{2} = 1 - \frac{\dot{h}_p}{r_p}t = 1 - \frac{\cos\xi\,\dot{x}}{r_p}t \tag{24}$$

If we assume that

$$\dot{x} = \frac{2\pi a_p}{P_{pl}} \tag{25}$$

then we can calculate the egress timescale as the time at which $\theta_p = 2\pi$, or

$$\tilde{\tau}_{eg} = \frac{P}{\pi a_p}\frac{r_p}{\cos\xi} \tag{26}$$

Therefore, the egress timescale allows us to constrain the value of $r_p/\cos\xi$.

5.2 Calculating Σ_s

To compute r_p from $\tilde{\delta}$, we must be able to compute Σ_s, which we can if we know (β_1, β_2) and ξ. If we define $\mathbf{i} = (i_1, i_2)$ and $\mathbf{f} = (f_1, f_2)$ as the points of intersection between the chord and the stellar disc, normalised to the stellar radius, then

$$i_1^2 + i_2^2 = f_1^2 + f_2^2 = 1 \tag{27}$$

The points of intersection must obey the vector equation

$$\mathbf{i} \quad \hat{\mathbf{c}} \quad \mathbf{f} \tag{28}$$

Where $\hat{\mathbf{c}} \equiv (\sin\xi, \cos\xi)$ is the unit vector with direction along the chord.

As the point $\boldsymbol{\beta} = (\beta_1, \beta_2)$ rests on the chord, we can write

$$\mathbf{i} + c_1\hat{\mathbf{c}} = \boldsymbol{\beta} \tag{29}$$

$$\boldsymbol{\beta} + c_2\hat{\mathbf{c}} = \mathbf{f} \tag{30}$$

where $c_1 + c_2 = c$. If ξ is known, we can solve for c_1 by rearranging:

$$\left|\boldsymbol{\beta} - c_1\hat{\mathbf{c}}\right| = \left|\mathbf{i}\right| = 1 \tag{31}$$

Which then allows the evaluation of $\mathbf{i}$. We can then make a similar calculation for c_2 and $\mathbf{f}$:

$$\left|\mathbf{i} + c_2\hat{\mathbf{c}}\right| = \left|\mathbf{f}\right| = 1 \tag{32}$$

And hence obtain $\mathbf{f}$, c_2 and the chord length c. With the chord length determined, this gives the central angle θ and ultimately Σ_s.

5.3 Breaking the Degeneracies

To recapitulate, the Shkadov transit has 3 measurements which are degenerate. The first is the transit ingress timescale τ_{ing}, which is degenerate in r_p and $b = \beta_2$; the second is the effective transit depth $\tilde{\delta}$, which is degenerate in r_p and Σ_s; the third is the Shkadov egress timescale $\tilde{\tau}_{eg}$, which is degenerate in r_p and ξ. Consequently, all three measurements are interdependent - measuring one constrains the values the others may take.

Assuming a value for ξ allows the observer to use $\tilde{\tau}_{eg}$ to obtain r_p, constraining $b = \beta_2$ from τ_{ing} and independent measurements of R_*. Measuring β_1 from the midpoint of Shkadov egress allows the calculation of Σ_s, which must be consistent with the transit depth $\tilde{\delta}$.

Therefore, model fitting of three measurements simultaneously can break the degeneracy. Bayesian methods are used commonly in parameter fitting of standard transit curves (e.g. [56, 57, 58, 38] and references within), so current techniques can be modified to accommodate the fitting of the Shkadov transit.

5.4 Masking the Shkadov Signal

Until now, we have considered idealised transit curves that are free of the sources of noise and error that real transit curves possess. This can include (but is not limited to) the underlying Poisson noise produced by photon processes, cosmic ray hits, scintillation effects produced by atmospheric turbulence, or other instrumental or calibration effects that (in combination) produce a background of time-correlated noise to all transit curves. Perhaps more dangerously, the presence of starspots can produce "dents" in the transit light curve, as the planet covers the starspot, which is by definition a cooler portion of the stellar disc (see e.g. [59, 60] for efforts to model this phenomenon). When an observer is trying to observe the duration of Shkadov egress, the presence of a starspot can make this extremely difficult. Figure 7 illustrates this problem. We have added time-correlated 1/*f* noise to the theoretical curve, produced by generating Gaussian white noise and applying the "pinkening" filter of [61]. To complicate matters further, we add a "starspot" to the curve, modelled by simply adding a Gaussian with a peak value slightly less than that of the planet to the curve. While not a particularly accurate starspot model, it demonstrates that unfortunately positioned noise sources can make the extraction of necessary Shkadov parameters particularly difficult.

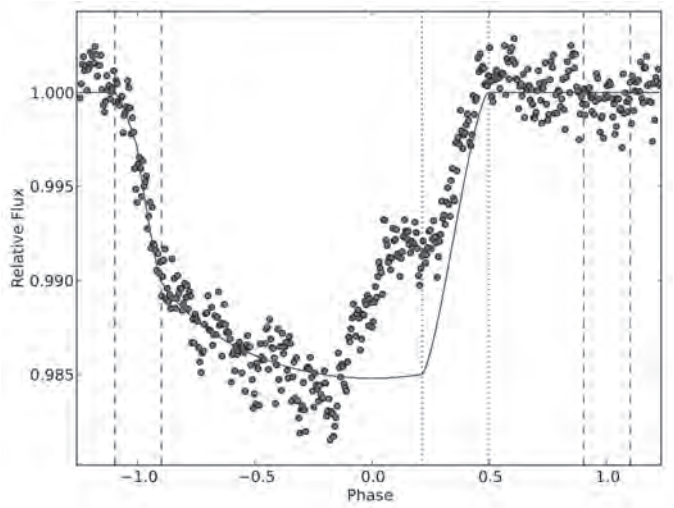


Fig. 7 The effect of noise on detecting Shkadov transits. The curve shows the analytically determined Shkadov transit curve, and the dashed and dotted lines refer to the ingress and egress points for normal and Shkadov transits, as in previous figures. The points represent the analytic curve with added 1/*f* noise and a starspot placed near the beginning of Shkadov egress.

6. DISCUSSION

6.1 Probability of Detecting a Shkadov Transit

Let the following events be defined:

- *T* - the system possesses a detectable transiting planet with a measurable curve
- *S* - the system possesses a Shkadov thruster
- *D* - the thruster is oriented such that its presence can be detected in the transit curve.

Given *T* and *S*, we can calculate the probability that the thruster will be detectable in the transit curve, $P(D|T, S)$. The thruster is a spherical segment of azimuthal angle 2ψ, and presumably of polar angle π (as this will maximise its thrust for any given ψ).

For the observer to see a thruster (with $\xi = 0^o$), then the position angle of the midpoint of the thruster α, defined along the line of sight of the observer (Fig. 8) must be in the range $[\pi/2-\psi, 3\pi/2+\psi]$.

Therefore

$$P(D|T,S) = P(\pi/2-\psi < \alpha < 3\pi/2+\psi) = \int_{\pi/2-\psi}^{3\pi/2+\psi} d\alpha' \quad (33)$$

which gives

$$P(D|T,S) = \frac{\pi+2\psi}{2\pi} \quad (34)$$

The maximum thrust achievable requires $\psi = \pi/2$, which would give a probability of effectively 1. In this case, there is only one possible configuration which does not give a Shkadov transit, where the mirror is positioned precisely behind the stellar disc, or equivalently $\alpha = 0$ exactly (although this particular configuration still allows for occultations of the planet by the Shkadov thruster - see following section). If $\psi = 0$, then

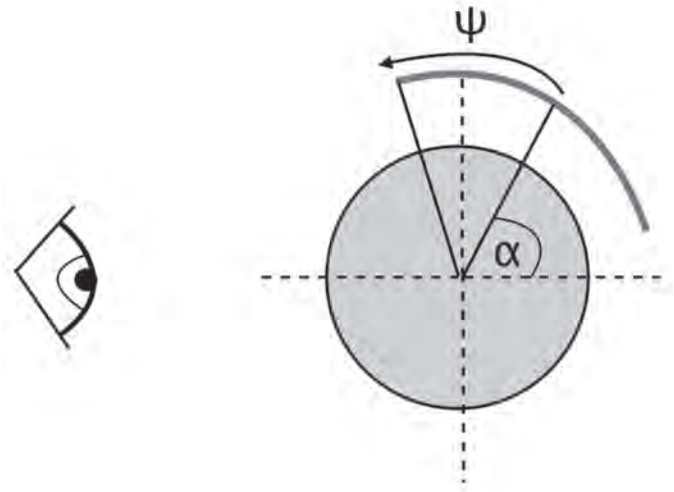


Fig. 8 Orientation of the thruster to the line of sight, as defined by the angle α.

the probability is 0.5 - the thruster is now an infinitely thin line across the stellar disc, which is either on the observer-facing side of the star or the other. In their calculations of thruster-assisted stellar motion, [51] use $\psi = 30^o = 0.523$ rad, giving a probability $P(D|T, S) = 2/3$. In the case where $\xi \neq 0^o$, we may expect that this probability will increase, as the total azimuthal extent of the thruster will be larger.

Of course, this conditional probability is less relevant than $P(D)$, which is given by

$$P(D) = P(D|T, S)P(S|T)P(T) \quad (35)$$

The probability that a system possesses a Shkadov thruster, given that the system possesses a transiting exoplanet, $P(S|T)$, is not well constrained at all. We do not know to sufficient confidence how many planetary systems possess intelligent life, and even less clear is how many of these intelligent beings would wish to build a Shkadov thruster. $P(T)$ is a combination of the probability that the star system has any planets at all, and the probability that the plane of a planet's orbit aligns

sufficiently well with the observer's line of sight, which for a circular orbit is

$$P_{align} = \frac{R_* + r_p}{a} \quad (36)$$

We can hazard a guess at $P(D)$ by appealing to parameter estimates for the Drake equation to obtain $P(S|T)$, assuming all stars have planets and P_{align} is of ~ 0.005 for an Earth orbiting a Sun-like star at 1 AU. If a fraction f_i of planetary systems possess intelligent life, a fraction f_c of intelligent species choose to communicate, and a fraction f_T of intelligent species decide to build Shkadov thrusters, then we can make an educated guess of

$$P(S) = P(S|T)P(T) \approx 10^{-3} f_i f_T \quad (37)$$

Given that recent surveys suggest that $f_i f_c < 0.01$ in the solar neighbourhood [3], and f_T is unconstrained, even for the very unlikely case of $f_T = 1$, the probability of detection is likely to remain around $10^{-4} - 10^{-5}$. More realistic values of $f_T < 0.01$ suggest that the probability of detection has an upper limit of $< 10^{-6}$. There are ~ 10^3 exoplanets detected to date, and only a fraction of the planet population in the vicinity of the Earth will be amenable to detection and characterisation via exoplanet transits, which suggests that the number of locally detectable Shkadov thrusters will be small. This reasoning is sufficient to rule out funding dedicated searches for Shkadov transits, but exoplanet transit surveys are likely to continue for the foreseeable future (see next section), and SETI scientists will have many opportunities to study public catalogues of data at low cost to attempt the detection of Class A stellar engines.

6.2 LIMITATIONS OF THE ANALYSIS

In this work, we have made several simplifying assumptions for the sake of more clearly illustrating the Shkadov transit as a potential detection method. In doing so, we have ignored aspects of the Shkadov transit curve that may prove important in future efforts to characterise it.

Perhaps most glaringly, we have ignored thermal emission from the planet in this analysis. By doing so, we discount the possibility of detecting secondary eclipses or occultations, as the planet passes behind the star. Depending on the orientation of the thruster, a secondary eclipse may be artificially lengthened if the thruster obscures the planet as it moves into occultation, or the mirror may reflect the night-side thermal emission back toward the observer, leading to a secondary eclipse curve that is "unbalanced": the thermal emission at the thruster side will appear enhanced compared to the non-thruster side. This may prove to be another means by which a Shkadov transit can be characterised.

We have assumed circular orbits in this analysis. The eccentricity of the planet's orbit plays an important role in characterising the transit curve. In particular, it can produce inequalities in the ingress/egress timescale of the transit. While in practice the difference amounts to less than a few percent [38, 62], highly eccentric orbits will confuse with the Shkadov signal. The extent to which this confusion will occur depends both on the eccentricity and the orbital longitude of periastron. Indeed, it may be the case that to fully characterise the Shkadov transit in this case, the secondary eclipse will be required to measure the difference between the impact parameter of the transit, b_{tra}, and the equivalent impact parameter for the occultation, b_{occ}, to constrain e.

Amongst our assumptions, we have required that the stellar radius R_* remains measurable despite the presence of the thruster. Stellar atmosphere modelling given external information on the star's properties remains the primary means of inferring R_*, but these models are constructed in the absence of technological interference with the stellar environs. The engine's presence artificially increases the effective temperature of the star. As the effective temperature is a key input variable in most stellar atmosphere models, there is a possibility that the radius will be misestimated. It is not immediately clear what radius the star would relax to after the thruster is placed, and it is possible that the star will present an increased oblateness depending on its rotation. In any case, using simple stellar atmosphere models may be insufficient (or at least inaccurate).

What is to be done instead? Interferometric measurements [63] or asteroseismology [55] may prove to be of assistance. Indeed, these measurements may be sufficient to confirm the existence of the Shkadov thruster through its effects on the star's structure. The unusual transit curve would merely serve as a flag to SETI scientists to require follow-up with one of these techniques!

It may even be the case that stars that do not possess transiting planets may still possess a Shkadov transit-esque signal, depending on the strength of their starspots. Indeed, the screening of one side of the stellar disc may set up a permanent Rossiter-McLaughlin effect, which would be detectable in the net spectral line profile of the star (cf. [64, 65]). While perhaps weaker than the signals produced in systems with exoplanet transits, future efforts to detect Shkadov thrusters, or projects which carry out photometric monitoring of stars, should be cognisant of this potential signal.

6.3 Detectability in Current and Future Surveys

A crucial factor in the ability to characterise Shkadov transits is the need for radial velocity follow-up, otherwise the semi-major axis of the planet's orbit will remain undetermined. Without this piece of information, the unmodified ingress timescale τ_{ing} cannot be used to produce the degenerate

$$k/\sqrt{1-b^2}$$

term, and the full deconstruction of the Shkadov transit is not possible.

If we are to identify current and future surveys that are likely to be able to yield fully characterised Shkadov transits, we should therefore focus on surveys that are amenable to radial velocity follow-up. The formidable success of the *Kepler* space telescope in detecting a large number of transiting exoplanet candidates [66] might lead us to think of it first, and indeed instruments such as HARPS-N have been designed with radial velocity follow-up of Kepler targets as a primary objective [67, 68]. However, parts of the Kepler sample will be too faint for current ground-based instrumentation to perform radial velocity follow-up [69].

Future surveys with a wide, shallow observing strategy such as TESS [70], or PLATO [71] are designed with radial velocity follow-up in mind. Selecting bright stellar candidates will allow

for better ground based measurements of the radial velocity, and provide a much larger sample of candidates to search for Shkadov transits. At the most extreme end, the GAIA satellite will provide a very large sample of low cadence measurements of stellar photometric and spectroscopic variations, providing a catalogue for follow-up missions to produce transit candidates [72].

The detection of a Shkadov transit is not contingent on high precision or sensitivity beyond being visible to RV instruments - although, depending on the Shkadov egress timescale, which is typically more abrupt than a standard transit egress, cadence may be a factor. If a candidate Shkadov transit is detected, missions designed to accurately characterise exoplanet transits to probe atmospheres, such as CHEOPS [73] or EChO [74] could be used to confirm that the curve does indeed show a Shkadov transit. In particular, these instruments can investigate the wavelength dependence of the Shkadov component. The Shkadov egress point may slightly shift with increasing wavelength due to diffraction, but we may expect the wavelength dependence of the unobscured transit ingress to be stronger.

Despite this, if the aim is merely to detect the Shkadov transit, and not to characterise its properties, then virtually any planetary system where a transit can be detected can yield evidence of a Shkadov transit. In this sense, searching for Shkadov transits can act as a means of flagging planetary systems for further SETI studies using other techniques.

7. CONCLUSIONS

We have outlined a means by which Class A stellar engines (or Shkadov thrusters) can be detected in exoplanet transit curves, if the stellar engine partially obscures the stellar disc during the transit. This constitutes a new serendipitous SETI detection method, which does not require the civilisation to be intentionally transmitting, and is present in the optical/infrared over a broad wavelength range. Also, as transit curves are periodic in nature, the signal is by definition repeatable.

We have shown the shape of the light curve in the presence of a Class A stellar engine, and demonstrated the deconstruction of the curve into the properties of both the transiting planet and the thruster (provided radial velocity follow up studies of the system are possible and the star's radius remains estimable). We note that both the primary and secondary eclipsesmay be affected, and that interferometric and asteroseismology measurements may help to characterise the nature of the stellar engine.

However, we also note that the detection of the features in the light curve produced by a Shkadov transit can be easily masked by phenomena such as starspots. As such, SETI scientists attempting to characterise Shkadov transits will require an excellent grasp of the sources of noise present in the light curve. Other measurements of stellar properties from radial velocity studies, interferometric imaging and asteroseismology may prove crucial in the full deconstruction of the properties of the stellar engine.

Exoplanet transit missions are of paramount importance in understanding the physical processes of planet formation, and are likely to remain so in the future. As such, while the probability of detecting a stellar engine is likely to be very low, current and future instruments will provide a large dataset from which to mine. We believe this will prove to be an extremely useful tool in constraining the population of Class A stellar engines in the Milky Way.

ACKNOWLEDGMENTS

The author gratefully acknowledges support from STFC grant ST/J001422/1, and thanks Caleb Scharf for useful comments.

REFERENCES

1. P. Morrison, J. Billingham and J. Wolfe, "The Search for Extraterrestrial Intelligence, SETI", in *The Search for Extraterrestrial Intelligence*, 1977.
2. J. Benford, G. Benford and D. Benford, "Messaging with cost-optimized interstellar beacons", *Astrobiology*, **10**, pp.475–90, 2010, doi: 10.1089/ast.2009.0393.
3. A.P.V. Siemion et al., "A 1.1-1.9 GHz SETI Survey of the Kepler Field: I. A Search for Narrow-Band Emission from Select Targets", *ApJ*, **767**, 94, 2013, doi: 10.1088/0004-637X/767/1/94.
4. H. Rampadarath, J.S. Morgan, S.J. Tingay and C.M. Trott, "The First Very Long Baseline Interferometirc SETI Experiment", *ApJ*, **144**, 38, 2012, doi: 10.1088/0004-6256/144/2/38.
5. A.J. Penny, "Seti with Ska", http://arxiv.org/abs/astro-ph/0408473. (Last Accessed 11 June 2013)
6. R.N. Schwartz and C.H. Townes, "Interstellar and Interplanetary Communication by Optical Masers", *Nature*, **190**, pp.205–208, 1961, doi: 10.1038/190205a0.
7. C.H. Townes, "At what wavelengths should we search for signals from extraterrestrial intelligence?", *PNAS*, **80**, pp.1147-1151, 1983.
8. A. Reines and G. Marcy, "Optical Search for Extraterrestrial Intelligence: A Spectroscopic Search for Laser Emission from Nearby Stars", *Publications of the Astronomical Society of the Pacific*, **114**, pp.416–426, 2002.
9. A.W. Howard *et al*., "Search for Nanosecond Optical Pulses from Nearby SolarType Stars", *ApJ*, **613**, pp.1270–1284, 2004, doi: 10.1086/423300.
10. J. Kraus, "We wait and wonder", *Cosmic Search*, **1**, pp.31–34, 1979.
11. R.H. Gray and S. Ellingsen, "A Search for Periodic Emissions at the Wow Locale", *ApJ*, **578**, pp.967–971, 2002, doi:10.1086/342646.
12. A. Hewish *et al*., "Observation of a Rapidly Pulsating Radio Source", *Nature*, **217**, pp.709–713, 1968, doi: 10.1038/217709a0.
13. A.J. Penny, "The SETI episode in the 1967 discovery of pulsars", *The European Physical Journal H*, 2013, doi: 10.1140/epjh/e2012-30052-6.
14. G. Benford, J. Benford and D. Benford, "Searching for cost-optimized interstellar beacons", *Astrobiology*, **10**, pp.491–498, 2010, doi: 10.1089/ast.2009.0394.
15. D.G. Messerschmitt, "End-to-end interstellar communication system design for power efficiency", http://arxiv.org/abs/1305.4684. (Last Accessed 11 June 2013)
16. A. Loeb and M. Zaldarriaga, "Eavesdropping on radio broadcasts from galactic civilizations with upcoming observatories for redshifted 21 cm radiation", *Journal of Cosmology and Astroparticle Physics*, **1**, 020, 2007, doi: 10.1088/1475-7516/2007/01/020.
17. D. Forgan and R. Nichol, "A failure of serendipity: the Square Kilometre Array will struggle to eavesdrop on human-like extraterrestrial intelligence", *International Journal of Astrobiology*, **10**, pp.77–81, 2010, doi: 10.1017/S1473550410000236.
18. S. Shostak, "Are transmissions to space dangerous?", *International Journal of Astrobiology*, **12**, pp.17–20, 2012, doi: 10.1017/S1473550412000274.
19. H.P. de Vladar, "The game of active search for extra-terrestrial intelligence: breaking the Great Silence", *International Journal of Astrobiology*, **12**, pp.53–62, 2012, doi: 10.1017/S1473550412000407.
20. J.E. Lovelock and I.R. Kaplan, "Thermodynamics and the Recognition of Alien Biospheres [and Discussion]", *Proceedings of the Royal Society B: Biological Sciences*, **189**, pp.167–181, 1975, doi: 10.1098/rspb.1975.0051.
21. N. Virgo, E. Simoncini and A. Kleidon, "The thermodynamic drivers of atmospheric chemical disequilibrium and their relationship to habitability", *EGU General Assembly 2012*, Vienna, Austria, 22-27 April 2012.
22. N.S. Kardashev, "Transmission of Information by Extraterrestrial

Civilizations", *Soviet Astronomy*, **8**, pp.217-221, 1964.
23. D.H. Forgan and M. Elvis, "Extrasolar asteroid mining as forensic evidence for extraterrestrial intelligence", *International Journal of Astrobiology*, **10**, pp.307–313, 2011, doi: 10.1017/S1473550411000127.
24. F.J. Dyson, "Search for Artificial Stellar Sources of Infrared Radiation", Science, **131**, pp.1667–1668, 1960, doi: 10.1126/science.131.3414.1667.
25. V.I. Slysh, "A search in the infrared to microwave for astroengineering activity", in *The search for extraterrestrial life: Recent developments*, **Vol. 112**, ed. M.D. Papagiannis, IAU, pp.315–319, 1985.
26. M.Y. Timofeev, N.S. Kardashev and V.G. Promyslov, "A search of the IRAS database for evidence of Dyson Spheres", *Acta Astronautica*, pp.655–659, 2000, doi: 10.1016/S0094-5765(00)00028-X.
27. R.A. Carrigan, "IRAS-Based Whole-Sky Upper Limit on Dyson Spheres", *ApJ*, **698**, pp.2075–2086, 2009, doi: 10.1088/0004-637X/698/2/2075.
28. M.M. Cirkovic, "Macroengineering in the Galactic Context: A New Agenda for Astrobiology", in *Macro-engineering: A Challenge for the Future*, eds. V. Badescu, R. B. Cathcart and R. D. Schuiling, Springer, pp.281–298, 2006.
29. J. Hickman, "The Political Economy of Very Large Space Projects", *Journal of Evolution and Technology*, **4**, 1999.
30. M.M. Cirkovic, "Against the Empire", *JBIS*, **61**, pp.246–254, 2008.
31. J.B. Campbell, "Archaeology and direct imaging of exoplanets", *Proceedings of the International Astronomical Union*, **1**, pp.247-250, 2006, doi: 10.1017/S1743921306009392.
32. D.P. Whitmire and D.P. Wright, "Nuclear waste spectrum as evidence of technological extraterrestrial civilizations", *Icarus*, **42**, pp.149–156, 1980, doi: 10.1016/0019-1035(80)90253-5.
33. J.S. Shklovsky and C. Sagan, "*Intelligent life in the universe*", Holden-Day, San Francisco, 1966.
34. R.A. Carrigan, "Starry Messages - Searching for Signatures of Interstellar Archaeology", *JBIS*, **63**, pp.90–103, 2010.
35. M. Mayor and D. Queloz, "A Jupiter-mass companion to a solar-type star", *Nature*, **378**, pp.355–359, 1995, doi: 10.1038/378355a0.
36. J.T. Wright and B.S. Gaudi, "Exoplanet Detection Methods", http://arxiv.org/abs/1210.2471. (Last Accessed 11 June 2013)
37. C. Lovis and D. Fischer, "Radial Velocity Techniques for Exoplanets", in *Exoplanets*, ed. S. Seager, University of Arizona Press, Tucson, pp.27-53, 2011.
38. J.N. Winn, "Exoplanet Transits and Occultations", in *Exoplanets*, ed. S. Seager, University of Arizona Press, Tucson, pp.55-57, 2011.
39. B.S. Gaudi, "Microlensing by Exoplanets", in *Exoplanets*, ed. S. Seager, University of Arizona Press, Tucson, 2011.
40. R.N. Bracewell, "Refinements to Drake's equation", *27th International Astronautical Congress*, Anahiem, California, 10-16 October 1976.
41. C. Walters, "Interstellar colonization: A new parameter for the Drake equation?", *Icarus*, **41**, pp.193–197, 1980, doi: 10.1016/0019-1035(80)90003-2.
42. C. Maccone, "A mathematical model for evolution and SETI", *Orig Life Evol Biosph.*, **41**, pp.609-619, 2011, doi: 10.1007/s11084-011-9260-3.
43. N. Glade, P. Ballet and O. Bastien, "A stochastic process approach of the drake equation parameters", *International Journal of Astrobiology*, **11**, pp.103-108, 2012, doi: 10.1017/S1473550411000413.
44. P. Hedelt *et al*., "Spectral features of Earth-like planets and their detectability at different orbital distances around F, G, and K-type stars", *Astronomy & Astrophysics*, **553**, A9, 2013, doi: 10.1051/0004-6361/201117723.
45. L.F.A. Arnold, "Transit LightCurve Signatures of Artificial Objects", *ApJ*, **627**, pp.534–539, 2005.
46. R. Conn Henry, S. Kilston and S. Shostak, "SETI in the Ecliptic Plane", Bulletin of the *American Astronomical Society*, **40**, p.104, 2008.
47. L. Arnold, "Transmitting signals over interstellar distances: Three approaches compared in the context of the Drake equation", *International Journal of Astrobiology*, in press, 2013.
48. L.M. Shkadov, "Possibility of controlling solar system motion in the Galaxy", 38th International Astronautical Congress, Brighton, United Kingdom, 1987.
49. D.R. Criswell, "Solar System Industrialization: Implications for InterstellarMigrations", in *Interstellar Migration and the Human Experience*, p.354, University of California Press, 1986.
50. M. Fogg, "Solar exchange as a means of ensuring the long-term habitability of the Earth", *Speculations in Science and Technology*, **12**, pp.153–157, 1989, .
51. V. Badescu and R.B. Cathcart, "Stellar Engines for Kardashev's Type II Civilisations", *JBIS*, **53**, pp.297–306, 2000.
52. A. Claret, "A new non-linear limb-darkening law for LTE stellar atmosphere models III", *A&A*, **428**, pp.1001–1005, 2004.
53. S. Seager and L. Hui, "Constraining the Rotation Rate of Transiting Extrasolar Planets by Oblateness Measurements", *ApJ*, **574**, pp.1004–1010, 2002, doi: 10.1086/340994.
54. J.W. Barnes and J.J. Fortney, "Measuring the Oblateness and Rotation of Transiting Extrasolar Giant Planets", *ApJ*, **588**, pp.545–556, 2003.
55. D. Stello *et al*., "Radius Determination of Solar-Type Stars using Asteroseismology: What to Expect from the Kepler Mission", *ApJ*, **700**, pp.1589-1602, 2009, doi: 10.1088/0004-637X/700/2/1589.
56. D. Veras and E.B. Ford, "Identifying Non-transiting Terrestrial Planets with Transit Timing Data", *Proceedings of the International Astronomical Union*, **4**, p.486, 2009, doi: 10.1017/S1743921308027002.
57. J.Z. Gazak *et al*., "Transit Analysis Package: An IDL Graphical User Interface for Exoplanet Transit Photometry", *Advances in Astronomy*, **2012**, pp.1–8, 2012, doi: 10.1155/2012/697967.
58. D.M. Kipping, "Luna: an algorithm for generating dynamic planet-moon transits", *MNRAS*, **416**, pp.689–709, 2011.
59. D.M. Kipping, "An analytic model for rotational modulations in the photometry of spotted stars", *MNRAS*, **427**, pp.2487–2511, 2012.
60. J. Tregloan-Reed, J. Southworth and C. Tappert, "Transits and starspots in the WASP-19 planetary system", *MNRAS*, **428**, pp.3671–3679, 2012.
61. N. Kasdin, "Discrete simulation of colored noise and stochastic processes and 1/f/sup a/power law noise generation", *Proceedings of the IEEE*, **83**, pp.802-827, 1995.
62. S.R. Kane, D.R. Ciardi, D.M. Gelino and K. von Braun, "The exoplanet eccentricity distribution from Kepler planet candidates", *MNRAS*, **425**, pp.757-762, 2012.
63. E.K. Baines *et al*., "Direct Measurement of the Radius and Density of the Transiting Exoplanet HD 189733b with the CHARA Array", *ApJ*, **661**, pp.L195–L198, 2007. doi:10.1086/519002.
64. Y. Ohta, A. Taruya and Y. Suto, "The RossiterMcLaughlin Effect and Analytic Radial Velocity Curves for Transiting Extrasolar Planetary Systems", *ApJ*, **622**, pp.1118–1135, 2005.
65. J.N. Winn *et al*., "Measurement of Spin Orbit Alignment in an Extrasolar Planetary System", *ApJ*, **631**, pp.1215–1226, 2005. doi:10.1086/432571.
66. N.M. Batalha *et al*., "Planetary Candidates Observed by Kepler. III. Analysis of the First 16 Months of Data", *ApJ*, **204**, p.24, 2013. doi:10.1088/0067-0049/204/2/24.
67. R. Cosentino *et al*., "Harps-N: the new planet hunter at TNG", in *Ground-based and Airborne Instrumentation for Astronomy IV*, eds. I. S. McLean, S. K. Ramsay, and H. Takami, **8446**, pp.84461V–84461V–20, 2012.
68. D.W. Latham and HARPS-N Collaboration, "*HARPS-N: A New Tool for Characterizing Kepler Planets*", American Astronomical Society, 2013.
69. P.J. Wheatley *et al*., "The Next Generation Transit Survey (NGTS)", eprint arXiv:1302.6592, 2013.
70. G.R. Ricker *et al*., "*Transiting Exoplanet Survey Satellite (TESS)*", American Astronomical Society, **42**, 2010.
71. C. Catala and T. Appourchaux, "PLATO : PLAnetary Transits and Oscillations of stars", *Journal of Physics: Conference Series*, **271**, p.012084, 2011.
72. M. Dennefeld, "The ESA-GAIA Mission: Not Only Astrometry", in *9th Pacific Rim Conference on Stellar Astrophysics*, eds. S. Qian, K. Leung, L. Zhu and S. Kwok, ASP Conference Series, **451**, 2012.
73. C. Broeg *et al*., "CHEOPS: A transit photometry mission for ESA's small mission programme", *EPJ Web of Conferences*, **47**, p.03005, 2013.
74. G. Tinetti *et al*., "EChO", *Experimental Astronomy*, **34**, pp.311–353, 2012.

(Received 23 May 2013; Accepted 30 May 2013)

* * *

ASTEROID CONTROL AND RESOURCE UTILIZATION

GRAHAM PATERSON[1], GIANMARCO RADICE[1]* AND J-PAU SANCHEZ[2]
1. *Space Advanced Research Team, School of Engineering, James Watt Building, University of Glasgow, Glasgow, G12 8QQ, UK.*
2. *Advanced Space Concepts Laboratory, Dept of Mechanical Engineering, University of Strathclyde, Glasgow, G4 0LT.*
Email: Gianmarco.Radice@glasgow.ac.uk**

Asteroids are materials rich small solar system bodies which are prime candidates for rendezvous and mining. Up until now much attention has been focused on methods of destroying or deflecting potentially hazardous asteroids from colliding with the Earth. Recently however the concept of asteroid capture has been suggested whereby the asteroid is returned to an orbit close to the Earth before mining can begin. This paper aims to provide a comprehensive introduction to the field for new researchers and to put forward a number of novel strategies for asteroid control.

Keywords: Asteroid control, asteroid resources, smart asteroid membranes, surface operations, asteroid mining

1. INTRODUCTION

There are severe problems facing humankind which will have to be tackled this century. Prominent among these are the scarcity of non-renewable resources [1] and global climate change [2]. It is possible that space utilization may provide significant partial answers to these problems. As one example solar power satellites could collect endless supplies of solar energy with no associated pollution cost before beaming the energy back to Earth [3]. To make this a reality space industrialization will be needed on a scale an order of magnitude higher than today's. This will demand a steep increase in construction resources having to be brought up the deep gravity well of the Earth, which at today's launch costs will be prohibitively expensive. A potential solution is to utilize the enormous resources that exist in the thousands of asteroids which orbit throughout the solar system. It is possible to rendezvous with a suitable asteroid and mine its resources, sending them back to Earth for exploitation [1]. Recently it has also been suggested that the asteroid orbit itself be altered, much like asteroid deflection scenarios for Earth defence, effectively shepherding it back to a position closer to the Earth [4]. Once in such an accessible orbit the asteroid can be more easily mined for its resources [5, 6]. The present paper will review the options available for making this possibility a reality.

2. ASTEROID LOCATIONS, COMPOSITION AND SHAPE

Asteroids are small rocky bodies, too small to be classed as planets or planetoids, orbiting the Sun. They are thought to be remnants of the early formation of the solar system. The three largest asteroids currently identified are Ceres (diameter 960 km), Pallas (570 km), and Vesta (530 km). The asteroids do not all orbit in the same location of the solar system, nor do they all have identical masses and compositions. The Main Asteroid Belt is between the orbits of Jupiter and Mars, between 2 AU and 3.5 AU from the Sun. More than 100,000 are known, with less than 5000 being larger than 10 km. Closer to the Earth there is a class of asteroids which holds the most promise of exploitation in the medium term, as well as potentially the largest threat to life on Earth itself – the Near Earth Asteroids [1]. These are defined to be asteroids with a perihelion q less than 1.3 AU. The proximity of these objects to Earth makes them obvious targets for rendezvous and exploitation.

The small size means that asteroid gravity fields are also exceedingly weak. Consequently, using gravity to remain attached to an asteroid is not feasible, as even slight reaction forces may cause detachment and escape. On the other hand the low gravity means that materials will be much lighter and also the intriguing prospect exists of using other forces to control surface dynamics, such as using electromagnetic forces to control orbits and influence the motion of dust clouds. Many space engineers have noted that the low gravity will almost certainly mean that a large asteroid bag will be needed to collect particles which have been knocked loose during the mining process [7].

There are also various classifications of asteroids according to composition. One of the most useful and prevalent in the literature is a subset of Tholen's taxonomy. One simple but widely used version of this characterises asteroids as C type, S type or M type. The C type carbonaceous asteroids are water rich and dark, making up the majority of asteroids. By contrast S type Silicon asteroids are based on iron, nickel and silicates, and are less common. Rarer still are M type metallic asteroids, which have high radar reflectivity based on a metallic nature.

According to current theories, heavy metals would sink to the core of a planet when it formed, a phenomenon known as differentiation, which is one reason geologists believe that large concentrations of iron and nickel exist close to the Earth's core. However on most asteroids it is likely that differentiation did not occur due to the low gravity, meaning that metal ores are likely to be more accessible in asteroids than on Earth. Studies of meteorites and spectral/photometric studies of asteroids have indicated that they are made of diverse but very valuable materials [1].

The surfaces of asteroids are extremely diverse, consisting

of boulders, craters, grooves and ponds. Craters are especially pronounced on Mathilde and Vesta, but appear on all the asteroids studied to date. Collisions between asteroids do occur occasionally, creating a pulverised ejecta blanket. This leads to a powdered regolith on larger asteroids, as they have enough gravity to attract some of the ejected material back. Smaller bodies may not have a regolith covering at all [8].

Asteroids are irregularly shaped, ranging from almost spherical for a few very large asteroids to potato shapes and even dumbbells. Their small size means that most asteroids do not have enough gravity to pull the surface to a sphere shape, and asteroid collisions ensure the irregular shapes after impact [8]. This has major implications for asteroid capture and mining because of the highly irregular nature of the gravity field around an asteroid.

3. ASTEROID RESOURCE POTENTIAL

The markets for asteroid resources can be divided into Earth based and space based [1, 9]. Earth based markets are based on non-renewable resources such as precious metals like platinum [9]. There are however difficulties involved, such as the expense of mining these asteroid resources, potentially lengthy periods of time before a return is made on an investment, and potential market instabilities due to large quantities of materials flooding terrestrial markets [6]. Space based markets are currently almost non-existent, but this situation can be predicted to change once large-scale space construction efforts gets underway. Specifically we may expect orbital research stations, orbital hotels for space tourism, space solar power satellites, spacecraft and assembly complexes, lunar bases and structures. Due to the deep gravity well of the Earth, building and supplying large structures in space is prohibitively expensive. However asteroids have a very shallow gravity well, and therefore asteroids have the potential to supply much of the materials for construction such as metals, as well as the fuel for spacecraft. Furthermore we may expect a market for unprocessed materials which may be used for radiation shielding in the harsh space environment. Hence we may expect a market for construction materials, semiconductors for fabrication of solar panels and electronics, volatiles for rocket fuel, water and oxygen for human habitats, unprocessed shielding materials [9]. If the asteroid was located at a linear Lagrange point the inherent instabilities around these points mean that very little fuel need be used to move the materials about the Earth-Moon system.

4. ASTEROID RENDEZVOUS AND ANCHORING

Ideally a target asteroid should have a semi-major axis close to the Earth's so that it is easier to reach, a low eccentricity to allow easier rendezvous, a low value of inclination due to the fact that plane changes are very expensive in terms of fuel, and initially it should be a small asteroid so that it is easier to control and shepherd back to Earth [1]. Following work by Shoemaker and Helin [10] it has been found that many asteroids are easier to reach than either the Moon or the Mars from an energy viewpoint. Note that the gravity force between the asteroid and the spacecraft is very small and for most asteroids extremely irregular due to the non-spherical shape. Consequently, stable 'orbits' around asteroids are not possible – however a series of small thruster firings for station keeping manoeuvres around the asteroid is possible. As a recent contribution, Kletetschka [11] has proposed a concept to use electromagnetic forces for station keeping manoeuvres around an asteroid with iron content. The concept would be to provide a magnetic field around the spacecraft, which would then either induce a magnetic field on the asteroid or interact with the permanent one already there. At spacecraft parking distances this field would be very weak but due to the very low nature of the gravity field it would not be insignificant, and it could theoretically be used to control motion around the asteroid. A magnetic coil could also lift up magnetic materials from the rubble pile regolith and collect them even without landing. Once orbit is acquired a suite of instruments can map the surface morphology and mineralogy, as well as study the dynamics of the asteroid in detail.

There are many options to anchor or tether to an asteroid. This can be done using harpoon penetrators/anchors, or alternatively a sling could be passed around the asteroid itself, indeed multiple slings have been proposed [7]. A sling looped around the asteroid spreads its load over a larger area than even a series of point anchors – slings are noteworthy for not violating the integrity of the asteroid. The advantage of multiple slings is that they spread the load even larger and are dynamically easier to control the asteroid. An alternative concept is the use of a net around the asteroid or indeed a membrane which is wrapped around the asteroid and then attached to the spacecraft. Another proposal is to pass a rope around the circumference, much like the sling, then tether to that rope. A combination of rope and attachment point anchors would bring us to a virtual asteroid monorail system with a 'monorail' structure along which equipment and astronauts could travel [7].

5. ASTEROID ORBIT CONTROL

Having accomplished rendezvous the next task will be to change its orbit, nudging it towards an easily accessible orbit for mining operations in near Earth space. Some authors have asserted that asteroids need to be brought under spin control before this can happen [12], while others have proposed schemes where spin control is not needed [13]. In either case the orbit can be modified by a propulsion system to move the asteroid towards a convenient easily accessible orbit closer to the Earth-Moon system. There are a variety of concepts for doing this, and we will briefly describe a few of these below.

5.1 Rocket and Sail Propulsion

It is possible to attach a rocket motor to an asteroid and use the thrust to change the orbit and send it back to Earth-Moon space. The major types of rocket we can use are a high impulse chemical motor, a low thrust ion engine, a solar thermal rocket, a nuclear motor, or a mass driver. Although mass drivers are theoretically possible to use, the propulsive effect comes from ejecting asteroid materials. In fact with some scenarios almost the entire asteroid would have to be consumed as reactive propellant. This is clearly not an attractive option. Even a chemical motor would require thousands of tons of fuel to make the required incremental changes in the asteroid orbit .We could propose high thrust nuclear engines or low thrust ion/plasma engines. However these require power sources which again have to be transported to the asteroids. A candidate for further research is a solar thermal rocket where In Situ Resource Utilisation provides water and focused solar energy from a mirror system heats the propellant. Finally we could attach a solar sail spacecraft to the asteroid and allow the pressure of sunlight to nudge the orbit gradually [5].

5.2 Enhanced Yarkovsky Effect & Focused Solar Energy

The Yarkovsky effect is a non-gravitational effect which results

from the anisotropic heating of an asteroid by the Sun [14]. On a rotating body which is exposed to the Sun the surface takes time to heat up and cool down and so the surface is warmer in late afternoon and dusk than in late night and dawn. The higher temperature surfaces radiate more thermal radiation, which leads to more heat being radiated thermally on the dusk side than the dawn side. The exceedingly small thrust levels and the complexity of the dynamics mean that Yarkovsky effects are not applicable to asteroid capture. However it has been suggested that enhanced Yarkovsky effects may be possible, where part of the asteroid is covered by an absorptive material such as charcoal or indeed painted black. This would increase the natural temperature difference and lead to an increases in net thrust. It is conceivable that using this technique would lead to feasible orbit alteration strategies. The use of a Smart Asteroid Membrane could enhance the Yarkovsky effect rather than paint or charcoal like materials. This membrane would have the ability to change its absorptivity in any desired manner from low to very high; much as modern materials can change their optical properties drastically in response to inputs such as electric fields. The use of a smart membrane would be advantageous because it could vary its absorptivity in response to commands or by on-board sensors, and the absorptivity could vary over the surface leading to potentially very rich dynamical effects which would be useful for orbital control and also attitude control [5]. Finally we can imagine focused solar energy as a propulsion system. This concept uses a mirror to focus solar energy onto the surface of an asteroid, sublimating part of the surface and using the reaction thrust to modify the orbit of the asteroid. Rather than one large mirror an alternative approach is to use a fleet of smaller mirror spacecraft to focus light on the asteroid. It may be envisaged that a formation of such spacecraft could deflect an asteroid onto an orbit in which it may be captured by the Earth- Moon system for parking and exploitation [15].

5.3 Field Forces-Use of Tractors

Several authors have proposed the use of field forces to nudge asteroids gradually into new orbits. Examples of these would be the use of gravitational, electrostatic and electromagnetic fields. The spacecraft are usually referred to as tractors in the literature. A gravity tractor is a spacecraft which uses gravitational attraction between the spacecraft and the asteroid to gradually nudge the asteroid into a different orbit. The spacecraft hovers close to the asteroid with thrusters angled outwards so as not to cause a thrust impingement on the surface. Using the mutual gravitational attraction between the spacecraft and the asteroid, it has been shown that a 20 tonne spacecraft can pull an asteroid gravitationally by hovering 25 m above the surface provided it can maintain a constant station-keeping thrust of 0.16 N [16]. A great advantage of the gravity tractor is that no actual attachments of any kind are necessary. The only force used is the natural force of gravity. Further there will not be any disruptive forces likely to fragment the asteroid en route back to Earth. A disadvantage is the need to angle the thrusters out to avoid thrusting onto the surface, so reducing the effective force. Another disadvantage is that the deflection force is exceedingly small and scales with mass, and so is smallest for very small asteroids. But it is precisely this population of asteroids which are likely to be targeted for the first missions to capture and exploit asteroids. Some form of increasing the attraction force would greatly increase the feasibility of the method.

Recently space engineers have suggested using electrostatic forces to control the orbits of asteroids, again in an asteroid deflection scenario for planetary protection [17]. This concept combines the elegance of field forces with the strength of electrostatic tractor beams, retaining the advantages of the gravity tractor and minimizing the disadvantages. The spacecraft and asteroid are both charged – the asteroid having been artificially charged by ions and electrons emitted from the spacecraft itself. This can result in the spacecraft and asteroid having opposite charge polarity or the same polarity. One of the problems with this approach is the uncertain nature of the asteroid surface. Major difficulties with electrostatic tractors are ensuring uniformity of the initial charge distribution on the asteroid surface, interactions with charges already on the surface, and the time evolution of the surface charge distribution. However these could be reduced and controlled by simply using a smart membrane and then charging the membrane rather than the uncertain surface of the asteroid. Again the use of a smart membrane advances the field due to replacing the uncertainties of asteroid surfaces with a well understood surface capable of fine control. It may again be worthwhile to examine in detail a new concept in which a smart membrane is used to get around these difficulties. The beauty of this novel idea is that it entirely removes the problems with the asteroid surface and composition, and not only allows very fine electrostatic control but also may allow the electrostatic properties to be minutely changed on the membrane in response to commands from the tractor.

6. RETURN TO EARTH-MOON SYSTEM

Returning the asteroid to the Earth-Moon system is feasible as we have seen [4]. A likely scenario is to aim the spacecraft/ asteroid into a stable manifold of the L1 or L2 regions, either of the Sun-Earth or the Earth-Moon system, allowing the asteroid to follow a Lissajous or Halo orbit around L1 or L2. It is of course feasible to use aerobraking to slow the asteroid down for capture. However public perception of this would likely be as a possible threat to earth and so this may not be feasible option, and of course there would also be a risk of fragmentation. Consideration would have to be given to the Roche limit of the Earth, particularly with the knowledge that many asteroids are weak rubble piles [18]. The Roche limit is defined to be the distance at which a body held together under its own gravity will be torn apart by the gravity field of a much larger body due to tidal forces. There is conventional wisdom to avoid bringing an incoming asteroid within the Roche Limit as it may be torn apart, however there may be a case for using the Roche Limit to break apart the asteroid deliberately. The reason for this is that it may then be processed more easily without drilling and conventional mining technologies. Of course a shattered asteroid will have to be enclosed in a very tough membrane to keep the materials together. This membrane must be loose enough to allow shattering to occur but strong enough to hold the shattered materials. This would of course require the development of very large and strong asteroid membranes.

7. ASTEROID MINING OPERATIONS

The first step in asteroid mining will be to ensure a proper means of anchoring equipment and structures to the asteroid has been achieved. This may use any concept previously discussed. Then suitable mining facilities and equipment will require being located on the asteroid. This will likely have been optimised already for asteroid mining and could be significantly different from terrestrial versions. There is a chance of the mined asteroid suffering a catastrophic failure and break up as the mass is removed and redistributed. To prevent this it is

possible to bind the asteroid together with rope, cable, mesh or membrane structures. A further possibility is to use a rigid case provided by a large processing facility. These asteroid restraint systems would also provide mechanical attachment points for processing equipment, habitats and so forth, and are virtually identical with the previously mentioned concepts of anchoring to the asteroid [7].

Starting with the concept of some systems of tethers/ cables (in fact any tension structures capable of attachment), it is possible to build mining platforms and machines which use the tension to provide reaction forces needed by drilling machinery. It is likely that mining will be concentrated on a small area initially. As mining goes deeper the need for a surface reaction force diminishes and tunnel forces become feasible. It is important that surface mining has a bagging element in place to catch the loose materials which are thrown up by the mining process itself. Diggers, material processors, power subsystems, habitats will be required. Of course the operation may be almost entirely robotic, although the stringent requirements of low gravity asteroid environments would likely require some form of human involvement, at least in its early stages [1, 7].

To mine ice composites an innovative approach is simply to enclose the asteroid in a membrane and then focus solar energy onto the asteroid. The water-ice would be sublimated and turned into mineral rich steam, which can be collected and processed, probably by standard fractional distillation techniques. Actually a portion of this water jet/ steam could be used for asteroid motion control, as previously mentioned. Eventually the asteroid will break up as its structural integrity diminishes, and so the bag must be strong enough to withstand the dynamic forces created when this happens. On the other hand, to mine rocky and metallic asteroids, the simplest technique is akin to surface strip mining on Earth, with the added complication of microgravity and the harsh space environment. Regolith can be shovelled, scraped and scooped. Indeed, simply scooping loose material off an asteroid may bring in astronomical profits [7, 9].

Transfer tugs will use conventional motors to move materials from the asteroid to LEO and back again. One of the main benefits of asteroids is that they are fuel stations for spacecraft. We can envisage a tug docking with an asteroid complex and refuelling using the processed volatiles from the asteroid itself. In concept, tugs are very well understood, and can be based on existing spacecraft such as the ATV. The likeliest scenarios are for tugs to transfer materials to an orbiting facility in LEO where the earthbound cargo can be transferred to rugged re-entry vehicles. The return vehicle can then return the materials to Earth either automatically or under manual control. The tug could then be refuelled and made ready for its next trip up to the asteroid processing facility. The other materials can then be used in LEO for building a space infrastructure or building solar panels. An alternative scenario would have the Earth return capsule separate automatically from the tug itself and re-enter without docking first at the LEO-receiving facility. Of course, the tugs may well transfer materials to the Moon, either directly or by a staged Lunar Orbital Station. This can then be used for construction, fuel, oxygen and water.

8. MANNED OPERATIONS ON ASTEROIDS

One of the great difficulties in exploiting asteroids is the lack of significant gravity. It is almost impossible to walk on the surface of a small body, due to very low gravity, non-vertical gravity vectors, variation of gravity over the surface, and the Coriolis forces on a spinning asteroid. To analyse walking physiologists use the Froude Number F, which is defined as $F = v^2/(g\ L)$ where v is walking speed, g is gravitational acceleration and L is a characteristic length, which we can take as leg length. It is found empirically that humans are walking optimally with F = 0.25 and start to move into a run at F = 0.5. On Earth this corresponds to a walking speed of about 1.5 m/s and a running speed of 2 m/s. On the Moon this drops to 0.6 m/s walking and running beginning at less than 1 m/s. Applying Froude to a typical small asteroid, Paterson calculated a walking speed of only 3 mm/s and a running speed of 5 mm/s [5]. It can be seen that walking on an asteroid is extremely difficult from a dynamic physiology point of view.

However there is another problem. The act of putting reaction force on an asteroid surface, necessary for movement, will have a high probability of making the astronaut ascend high above the surface, perhaps into orbit or escape the gravity field of the asteroid completely. Further, astronauts falling to the surface would do so at a much slower rate than happens on Earth. This may be a positive result in certain scenarios. For example an astronaut could perform work and still have a long time before impacting the surface. Conversely, an astronaut who found themselves suspended for whatever reason would have no way of quickly reaching the surface again until the weak gravity had pulled them down. Again, on the positive side, tools may be left virtually suspended beside the astronaut which may be advantageous but loose chippings about the astronaut will take a long period to settle again – perhaps surrounding a miner with a semi-permanent cloud of dust. Alternatively they may go into orbit, or indeed leave the asteroid altogether if above the escape velocity. The asteroid gravity again turns out to have positive and negative consequences [5].

The problems highlighted have resulted in some studies on systems for the personal motion of astronauts over asteroids. For example, the system pioneered at MIT is the Aster Rope [20].In this system two ropes are secured around the asteroid and the tension forces used to hold down astronauts and machinery. Each rope is threaded through eyelets connected to a pair of bars on each side of the astronaut, which are then in turn connected to a gimballed waist ring permitting rotation along three different axes. It has been found by MIT researchers that the best available material for this system is Vectran, which is space qualified. The astronaut does not actually have to walk along vertically of course – they could just as easily move along horizontally, perhaps with a motor attached to the rope for propulsion. Another option is a free flyer of some sort. This is similar to the Manned Manoeuvring Unit of the Shuttle but optimised for use in low gravity environments such as asteroids. A space miner could have a free flyer optimised for travel over the asteroid surface. It could have propulsion, power, fuel, manipulators and a suite of attached tools such as drills which have been optimised for use on asteroids [5]. A strong tether may be attached to a surface point for extra protection in case of an accident. It could have a small storage area for interesting samples, and perhaps even a small nano-laboratory for on the spot analysis. It is possible to envisage a small built in remote sensing unit which can provide an analysis of the surface/ near subsurface, perhaps with an attached ‘mini-harpoon’ for subsurface analysis. It is possible to envisage electromagnetic forces for station keeping for some classes of asteroid. Finally, the use of underground tunnels is attractive for many reasons. Of course the initial work will have to be on the surface but as soon as tunnelling commences the astronaut miners will have a means not only of accessing deep veins of minerals but also

of using the reaction forces within the tunnels for stabilisation and drilling. The tunnels would provide protection against space radiations as an added bonus. We could even imagine the asteroid being progressively mined and hollowed out, with large areas becoming available for human habitation – it is feasible to imagine even environmental control being possible such as pumping air into the tunnels for breathing and a system of heat management units. However there are still multiple problems that will need to be solved [7].

Clearly strategies for moving over the surface of an asteroid, tools applicable to this environment, artificial manned manoeuvring systems which can function in a low gravity environment, a personal tether systems of some sort, personal motion systems to allow walking in this environment, systems for dealing with persistent dust clouds and debris flying off work areas, retrieval systems for tools inadvertently rising too quickly from the work site, all will require to be researched and engineered ahead of actual manned mining operations on an asteroid.

9. POLITICAL ISSUES

The 1967 Outer Space Treaty was the first international space treaty, and still the most important [21]. For the purposes of asteroid engineering the most important sections relate to ownership by private companies. Article II of the treaty states that "Outer space, including the moon and other celestial bodies, is not subject to national appropriation by claim of sovereignty, by means of use or occupation, or by any other means." Theoretically, this prohibits only national and not private appropriation of territory. However, in common law countries such as the United States, legal theory dictates that the government must have sovereignty over territory before it can confer title on its citizens. Consequently, traditional real property rights would not be possible under this regime. According to retired space law Professor Frank Lyall formerly of Aberdeen University while it is not legal to own an asteroid, it may be different if someone simply knocks a chuck off it and then brings it back [22]. Further, liability laws for space operations will mean that the entire asteroid engineering lifecycle will require extensive insurance, including liability cover in case the return to Earth orbit and Earth's surface goes wrong.

We must be aware of another issue. This is the issue of a real or perceived accident in mining an asteroid which leads to a large fragment or fragment shower impacting the Earth or even the Moon. How likely is such an event? The reality and especially the perception must be that it is virtually impossible for an Earth threatening disaster to occur. This applies to the Earth approach phase as well as the parking phase of operations, and even applies to contamination issues for materials brought back from the asteroid.

10. CONCLUSIONS AND RECOMMENDATIONS

This paper has demonstrated that capturing an asteroid and returning it to Earth for processing is very difficult but not beyond our reach. It is also clear however that the asteroids contain an enormous untapped resource for mankind in areas ranging from Earth based precious metal markets to space based construction, rocket fuel and system support. It may be argued strongly that if asteroid mining does not take place mankind's expansion into the Solar System will be long, difficult and of minimal effect. Rendezvous with the most accessible asteroids is possible using conventional systems, and it is possible to attach a variety of novel propulsion units to an asteroid to bring it back to the Lagrange corridors of Earth, where it can be put into a parking orbit. It is possible to mine these asteroids if an effort is made to develop low gravity mining systems and return the mined materials to the Moon for construction, to Earth for commerce and industry, and to LEO for manufacture. The materials returned to Earth have enormous monetary values, and industrial investors will have their initial investment repaid back many times, which will stimulate further capture schemes. The exploitation of asteroids has the potential to catalyse the coming together of nations as partners and move polluting heavy industry off Earth in a phased and stable evolutionary process. It is possible to consider questions of safety and build these into the initial planning phase to make sure that no captured asteroid can possibly threaten the Earth. Much further research needs to be done on all these issues to ensure that humankind does indeed develop an optimal technology to capture and exploit the enormous resources which have been placed tantalisingly close to us, which will stretch us to reach, but which will allow us to expand into the solar system and become a true solar system wide civilization.

REFERENCES

1. S.D. Ross, "Near earth asteroid mining", www2.esm.vt.edu/~sdross/papers/ross-asteroid-mining-2001.pdf. (Last Accessed 29 January 2013)
2. R.K. Pachauri and A. Reisinger (eds), "Climate change 2007 synthesis report", www.ipcc.ch/publications_and_data/ar4/syr/en/contents.html. (Last Accesed 29 January 2013)
3. A. Martin (ed), "Solar power satellites and spaceplanes", www.reactionengines.co.uk/tech_docs/ssp_skylon_ver2.pdf. (last Accessed 29 January 2013)
4. D. Massonnet and B. Meyssignac, "A captured asteroid: our David's stone for shielding earth and providing the cheapest extra-terrestrial material", *Acta Astronautica*, **59**, pp.77-83, 2006.
5. G. Paterson, "Asteroid capture and exploitation", MSc Thesis, University of Glasgow, Glasgow, UK, 2011.
6. M. Sonter, "The technical and economic feasibility of mining the asteroids", *Acta Astronautica*, **41**, pp.637-647, 1997.
7. J. Remo (ed), "*Near Earth Objects: The UN International Conference*", **Vol 882**, New York Academy of Sciences, New York, 1997.
8. NASA Lunar and Planetary Sciences, Asteroid Homepage http://nssdc.gsfc.nasa.gov/planetary/planets/asteroidpage.html. (Last Accessed 19 January 2013)
9. C. Gerlach, "Profitably Exploiting Near earth Object Resources", *International Space Development Conference 2005*, Washington, DC, 19-22 May 2005.
10. E. Schoemaker and E. Helin, "Earth Approaching Asteroids As Targets For Exploration", in *Asteroids: an Exploration Assessment*, Morrison and Wells, 1978.
11. G. Kletetschka, T. Adachi and V. Mikula, "Electromagnetic spacecraft used for magnetic navigation and levitation within asteroid belt, mining concepts and asteroid magnetic classification", *38th Lunar and Planetary Science Conference*, League City, Texas, 12-16 March 2007.
12. D. Scheeres and R. Schweickart, "The Mechanics of Moving Asteroids", *Planetary Defense Conference, Orange County*, 23-26 February 2004. AIAA paper 2004-1446.
13. J. Sanchez and C. McInnes, "Assessment on the Feasibility of Future Shepherding of Asteroid Resources", *61st International Astronautical Congress*, Prague, Czech Republic, 27 September - 1 October 2007. Paper No, IAC-10.D3.2.7.
14. D. Vokrouhlicky, S. Chesley and A. Milani, "On the observability of radiation forces acting on near earth asteroids", *Celestial Mechanics and Dynamical Astronomy*, **81**, pp.149-165, 2001.
15. M. Vasile and C. Maddock, "On the deflection of asteroids with mirrors", *Celestial Mechanics and Dynamical Astronomy*, **107**, pp.265-284, 2010.

16. E. Lu and S. Love, "Gravitational Tractor for towing asteroids", *Nature*, **438**, pp.177-178, 2005.
17. N. Murdoch and D. Izzo, "Electrostatic Tractors for Near Earth Object Deflection", *59th International Astronautical Congress*, Glasgow, UK, 29 September - 3 October 2008. Paper No. IAC-08-A.3.1.5.
18. N. Movshovitz and E.Asphaug, "Long term stability of a rubble pile Phobos", EPSC Abstracts, **Vol. 6**, EPSC-DPS2011-1654-1, 2011.
19. A.E. Minetti, "Walking on other planets", *Nature*, **409**, pp.467–468, 2001.
20. I. Garrick-Bethell and C. Carr, "Working and walking on small asteroids with circumferential ropes", *Acta Astronautica*, **61**, pp.1130-1135, 2007.
21. "*UN Treaties on Outer Space*", UN Publications, 2002.
22. F. Lyall, private communications with one of the authors (GP), 2010.

(Received 21 November 2011; Accepted 14 November 2012)

* * *

JBIS, Vol. 66, pp.161-166, 2013

APPLICATION OF COTS COMPONENTS FOR MARTIAN SURFACE EXPLORATION

MATTHEW CROSS, CHRISTOPHER NICOL, ALA' QADI AND ALEX ELLERY
Department of Mechanical and Aerospace Engineering, Carleton University, 1125 Colonel By Drive, Ottawa, K1S5B6, Canada.

This paper details the motivation behind COTS component testing, the results, and their application for a low-cost, sample and fetch scout rover to accompany large scientific exploration rovers. Carleton University's Space Exploration Engineering Group proposed using Maxon motors and controllers for the wheel and mechanism assemblies on a micro-rover development platform called *Kapvik*. These components were subjected to extensive environment tests to determine their operability at sub-rated temperatures. The results showed the motors and controllers will perform at temperatures experienced at equatorial Mars.

Keywords: Mars rovers, COTS components

1. INTRODUCTION

Kapvik is a 30 kg micro-rover analogue designed as a tool for further developing Canada's planetary exploration capabilities. *Kapvik* was funded by the Exploration Surface Mobility (ESM) project of the Canadian Space Agency (CSA) and administered by MPB Communications Inc. This prototype is designed with a view to flight qualification and to help asess potential exploration missions to which Canada may contribute. It was designed for temperatures associated with summer in the high arctic. Its trial operations will be in an unknown environment - likely in the Canadian Arctic - analogous to the Martian equatorial surface.

The Space Exploration Engineering Group (SEEG) at Carleton University was responsible for the design and development of the chassis, motor control, and additional instrumentation mechanisms for *Kapvik*. *Kapvik*'s chassis comprises of six wheels in a rocker-boogie system each individually driven by a Maxon motor. The rover must be able to operate in temperatures of -20 ºC to 40ºC for evaluation purposes. In addition, the entire rover was designed in a way that provides a "path to flight", i.e. it uses components that have a suitable space qualified counterpart, or demonstrate suitability for space application. The chosen motor controllers have a rated operating temperature between -10ºC and 45ºC. This is a narrow window when compared to the overall required temperature range for the rover. *Kapvik* has a limited power budget of 30 W due to the constraints of being a micro-rover. This is the motivation behind a passive thermal control system. The motor and controller pair's performance was characterized in an effort to determine how tolerant the motor control would be at sub-rated temperatures.

2. THE *KAPVIK* ROVER

Kapvik (Fig. 1) has an instrumented six-wheeled rocker-bogie system with differential drive similar to NASA's fleet of exploration rovers: *Sojourner* [1], *Spirit* [2] and *Opportunity* [3], and *Curiosity* [4]. The rocker-bogie allows all six wheels to maintain ground contact to enhance mobility while allowing the rover to climb over rocks [5, 6]. Each of the six rigid wheels have 24 grousers. The grousers extend 5 mm from the base of the 145 mm diameter wheel giving the wheel an average diameter of 150 mm. *Kapvik* contains a suite of instruments for localization including a laser range finder, stereo vision camera, inertial measurement unit (IMU) and sun sensor. A robotic arm comes equipped with a scoop for collecting soil samples. Solar panels are mounted on top of the cab to provide unregulated power at 24 V. Inside the cab, the avionics system includes the motor controllers, on-board computing and power electronics. When fully operational, *Kapvik* is capable of semi-autonomous navigation using localization and path planning.

Nomenclature

CAN	Controller Area Network
COTS	Commercially Available Off-The-Shelf
CSA	Canadian Space Agency
ESM	Exploration Surface Mobility
IMU	Inertial Measurement Unit
MER	Mars Exploration Rover
MFEX	Pathfinder Microrover Flight Experiment
SEEG	Space Exploration Engineering Group

2.1 Instrumented Mobility System

An instrumented chassis and individually throttled motors were early design objectives for *Kapvik*'s mobility system. The instrumented chassis allows for the mobility system to adapt to changing terrain conditions. Each motor can be individually throttled to allow for traction control on Martian rovers [7, 8].

Each of the six wheels is driven by a Maxon Motor RE25 motor with planetary gearing and harmonic drive. The wheel motor assemblies include planetary gearing and a harmonic drive for a gear ratio of 1400:1, and are contained at the wheel base inside an enclosure to protect against weathering and dust. Wheel odometry is provided by incremental encoders attached to each motor. Single axis load cells are mounted on top of each wheel to measure the vertical force. Potentiometers provide the

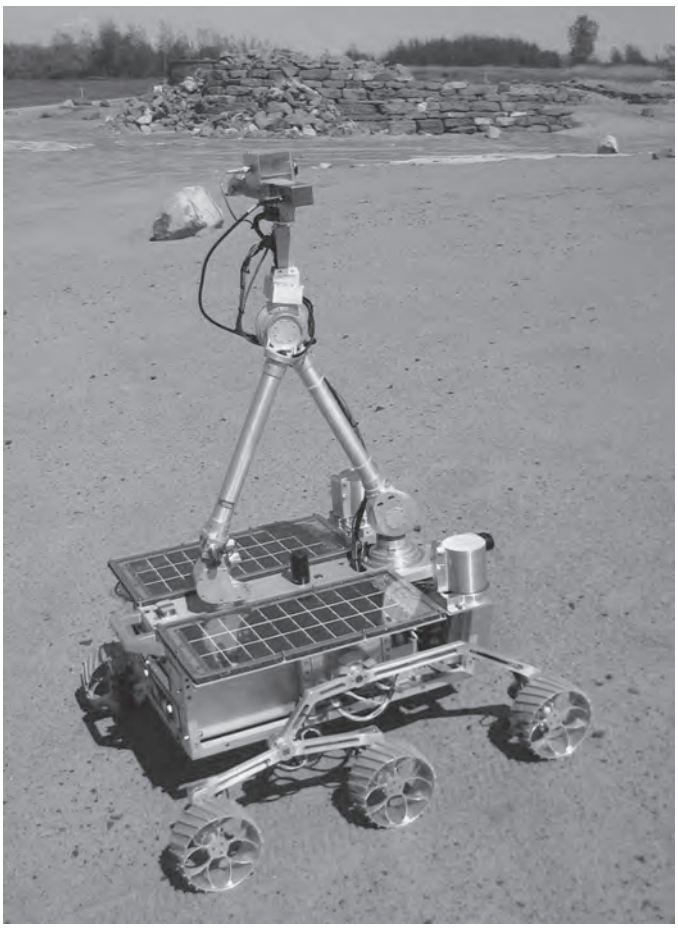


Fig. 1 ***Kapvik's*** **mobility system, including the motor controller interface, was developed by the SEEG team at Carleton University.**

rotational angles of the differential drive and the rockers. When calibrated, the potentiometers provide an orientation of each of the wheels relative to the body frame of the rover.

Each of the motor assemblies is powered and driven by a Maxon Motors EPOS 24/1 motor controller. The potentiometers, load cells and incremental encoders all connect to the motor controller; all of the motor controllers are connected via controller area network (CAN) bus to a central computer on a Xiphos Q6 card, which is located within the avionics enclosure on the rover. The controllers are capable of three primary control modes: constant current, constant velocity, and fine position control. The constant velocity mode is used to keep each all of the *Kapvik* wheels driven at the same rotational velocity. This PI control mode samples the encoder and drawn current data at 1 kHz to maintain a fixed rotational wheel velocity. The fine position control mode is similar in that the controller maintains the motor at a constant wheel velocity. The difference is the control scheme takes into account initial position and final position. This PID control mode samples the encoder and drawn current data at 1 kHz and is used for fine position control of the vision system and laser scanner position mechanisms. The fixed current mode is not presently used for any control mode for *Kapvik* but is included for completion.

2.2 Comparison to NASA Sojourner

The NASA *Sojourner* rover, previously known as the Pathfinder Microrover Flight Experiment (MFEX), faced similar challenges in its development. The MFEX team of engineers identified several components, including the motors and gears, as technological challenges to operating a low-cost micro-rover at Martian surface temperatures down to -80ºC and survival down to -110ºC. Prior to MFEX, brushless motors were the preferred option for space applications. However, their integrated electronics were deemed unsuitable for the Martian temperatures and moving the electronics inside the avionics box was deemed unsuitable for the amount of wiring that would be crossing over moving parts. The MFEX engineers instead selected brushed Maxon motors for their torque-to-mass performance. While the motors and gears were selected off-the-shelf, MFEX engineers worked with Maxon to modify the units to optimize them for the low temperatures. The capacitors were encapsulated to improve power usage, and the gearing was suitably lubricated for the temperature. The resulting actuator was shown to operate without degradation over 8 km [9].

For *Kapvik*, it was not appropriate to have any custom modifications at its current development state: it is not a flight model. The entire mobility system was designed with the "Surrey Approach"; that is, it was designed with commercially available off-the-shelf (COTS) components and tested to meet its environmental criteria. Most off-the-shelf components, such as the motors, have space qualified versions or modifications available. Customized parts, such as the rocker-bogie, were designed to be space qualifiable.

2.3 Thermal Requirements and Component Selection

CSA requires *Kapvik* to be stored in temperatures from -40ºC to +50ºC and operate at -20ºC to +40ºC. The lower bound of temperatures corresponds to expected conditions in Northern Canada and equatorial Mars. Table 1 summarizes the thermal requirements for the *Kapvik* micro-rover as well as the rated performance of the selected mobility equipment.

The Maxon motors were selected for their heritage on previous Mars rovers. The Maxon RE25 motor is not space qualified however it best met the operational requirements for *Kapvik*. The Maxon EPOS 24/1 controllers were selected for their low mass, small form factor, and compatibility with the motors. The use of the EPOS controllers with Maxon motors is fully developed and documented, eliminating any interface development between the motors and controllers. All motors are controlled using the same type of controller which further reduced the software development time. Since each controller is a self contained unit they satisfied another requirement, design modularity. Additional motor controllers could be added as additional nodes on the CAN Bus. The Maxon motors rated down to -30ºC; however, the EPOS 24/1 controllers are rated to only -10ºC.

The motivation for subjecting the motor and controller assemblies to temperature tests was to increase the confidence in their performance at sub-rated temperatures. That way the off-the-shelf components could be used without modification and there would be no need to actively control their environment. An actively controlled temperature would increase the complexity of the motor assembly enclosure and increase the power demand on an already limited power budget. The motor controllers are located within an avionics enclosure that would be at a higher temperature than the motors during operation; however, the motor controllers were at this point deemed to be critical component to test for compliance. It was only found after the test data analysis that the rated temperature for the harmonic drive was only 0ºC.

TABLE 1: *Operating and Storage Temperatures (ºC) for Selected Components.*

	Operation		Storage	
	Minimum	**Maximum**	**Minimum**	**Maximum**
Mission Requirements	-20	40	-40	50
Maxon RE25 Motor	-30	100	-30	100
Maxon planetary gear	-40	100	-40	100
Maxon EPOS 24/1 Controller	-10	45	-40	85
Encoders	-25	85	-25	85
Harmonic Drive	0	50	0	50

3. TEMPERATURE TESTS

The temperature experiments were conducted in an environmental test chamber (Fig. 2) and housed in the Department of Electronics at Carleton University. The component is rated to be stored in temperatures as low as -40ºC and testing was done as cold as -30ºC. Three motor and controller pairs were connected by CAN Bus. One pair was kept in the environment chamber, and the other two pairs had one motor or controller inside and the other outside. The three pairs ran simultaneously for temperatures ranging from +40ºC down to -30ºC at 5 degree intervals. At each temperature, each of the three operating mode profiles was run three times. The chamber was set to a desired temperature and the components were allowed to sit for many hours to allow the entire set-up to reach a steady state.

The first test was to determine if the controller would pass a desired current to the motors in sub-rated temperatures. Since this is not a binary pass or fail, the metric chosen to judge the trend of the performance over temperature was the standard deviation between the input and the output current and velocity for different modes. In order to determine if the performance was adequate it must be compared to the rest of the range of temperatures the component will experience. The effects of temperature change of each the motor and controller pair were separated and then the combined effect was viewed. Tests were done using constant velocity mode, constant current mode and position mode. Velocity mode was the nominal control mode for the wheel motors. If the results showed the performance of the velocity mode deteriorated much more rapidly in sub-rated temperatures than the current mode, it may be worthwhile to take advantage of this mode in some way. Early tests showed that without a start-up phase, a cold motor would produce very noisy results for the first number of runs. This was attributed to the fact that lubrication in the motor and gearing has more viscous effects when cold. Subsequent experimental runs had a warm-up phase with current applied to the motors to warm up the lubrication before the actual test was done.

At the time of testing, *Kapvik*'s motor control software was developed enough to be incorporated into these temperature experiments. At colder temperatures the motors were allowed to run for a warm-up period before running a series of constant speed, constant current, and positioning profiles. For the constant current mode, the demand current, measured current drawn and measured output shaft speed were sampled at 10 Hz and recorded for later analysis. The resolution of the current monitor is 1 mA. Similarly for the constant speed mode, the mode at which the wheels are operated, the demand speed, measured output speed, and measured current drawn were sampled and recorded. The resolution of the encoders are 1 RPM. The position mode was similar to the velocity mode with the only difference being the demand speed was the pre-defined position profile speed. For each of the operating modes, the magnitudes and directions of the demand values were altered to show the step response. The duration varied from approximately 1 minute at each step down to only a few seconds.

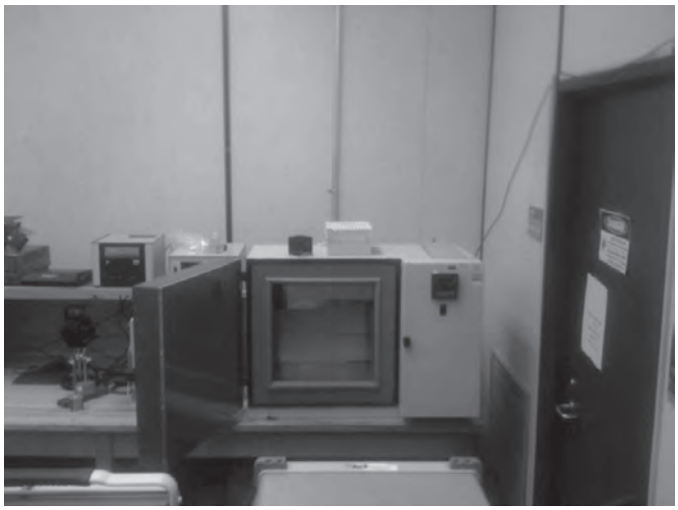


Fig. 2 Environmental Chamber used for testing at Carleton University.

4. ANALYSIS OF RESULTS

The initial purpose of these tests was to determine whether or not the controllers could operate at -20ºC. Figure 3 shows the standard deviation of the error for the three nodes tested over the range of temperatures from -30ºC to +40ºC when controlled with the velocity mode. These curves show that there is similar performance in velocity mode regardless of the controller's temperature with the motor at a constant temperature. This means that the controller is able to maintain the desired output velocity to some accuracy across the temperature range. As the motor temperature decreases, the standard deviation remains consistent until -20ºC and begins to increase below that. This would once again indicate the controller is able maintain the desired output velocity across the required temperature range.

Looking at the current output of the controller, there is a bias in performance depending on the direction the motor spins. The results show that the current draw decreases at lower temperature in the positive current direction and increases in the negative current direction. This divergence is likely due to the warm up phase where the motor is spun in the positive current direction for a few minutes. The difference in current

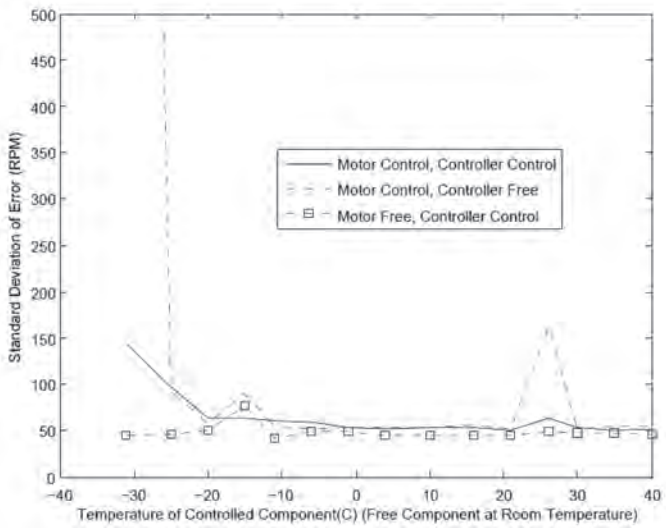


Fig. 3 Results of the motor and controller in velocity mode for different temperatures.

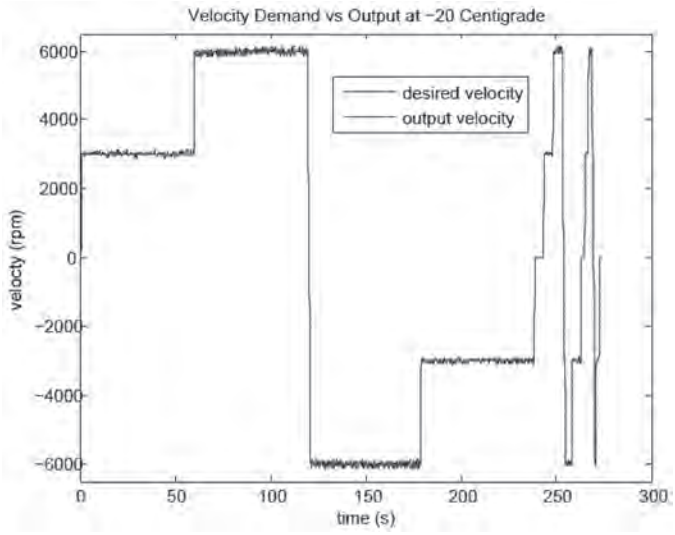


Fig. 4 Desired and output velocity with cold motor at -20ºC.

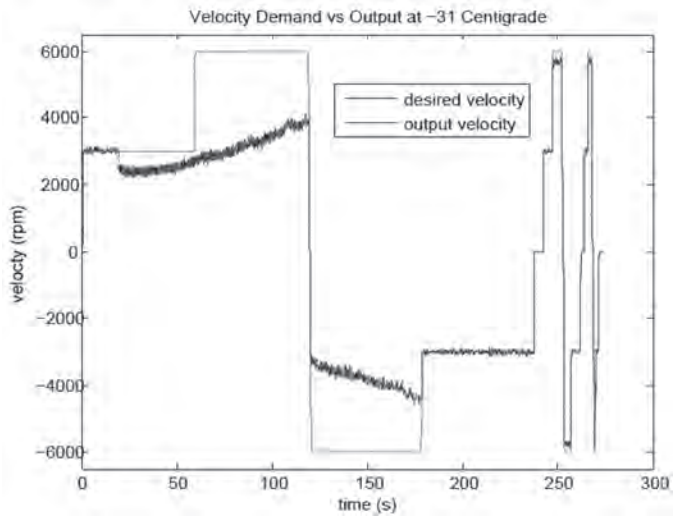


Fig. 5 Desired and output velocity with cold motor at -31ºC.

drawn between room temperature and -20ºC is +/- 25 mA depending on the direction. This difference increases up to +/- 40 mA at -30ºC. The current draw is, as expected, greater at lower temperatures. To meet the same velocity profile the current requirement increases four-fold at -20ºC.

Figures 4 and 5 show the velocity output along with the desired velocity profile used during testing when the motor is at -20ºC and -31ºC respectively. In both cases the motor controller is at an ambient room temperature. Despite no change in the motor controller's temperature, the decrease in the motor temperature causes the velocity output to be unable to match the desired profile. This shows that it is not the motor controller, but the gearing that is the critical component in the chassis design.

Table 1 shows that while the motor is rated to -30ºC, the harmonic drive attached to the motor is only rated down to 0ºC. Schafer [10] identified the critical interface to be between the wave generator and the flexspline within the harmonic drive. As the temperature decreases, the lubrication becomes more viscous. The motor has to work harder to overcome this greater friction. Figures 6 and 7 show the change in both efficiency and the no-load-torque respectively with temperature. This additional force/loss of efficiency could be the cause of the problems seen at -31ºC.

The constant current mode, while not used in *Kapvik*'s operation, showed very little noise at any temperature. These results can be seen in Fig. 8. However, the output velocity greatly decreases. A simple calculated rate of current drawn per output speed shows the drastic difference in performance for the two modes. At room temperature, this ratio is similar for the two modes. However at -20ºC the ratio is 5 times greater in current mode than velocity mode. Furthermore, the ratio in velocity mode is 5 times greater at -20ºC than it is at room temperature. It is better to use a slightly noisy velocity control than a more precise current control as it is more energy efficient.

5. COTS SOLUTION FOR MARTIAN EXPLORATION

The preliminary thermal tests show it is possible to utilize commercially available motors and controllers for applications down to -30ºC without modifications. The thermal requirements for an analogue mission have been met; however, this interest lies in the path to flight for a micro-rover using off-the-shelf hardware. In this section we shall outline how this applies to a fetch and retrieve rover for future Mars surface such as with *ExoMars*.

Mars surface exploration is expensive. Academic and agency scientists spend years carefully developing scientific return case studies. Scientific instruments then are engineered to study Mars soil, rocks and atmosphere. Large, sophisticated rovers take years to develop. After several years and billions of dollars, Mars exploration rovers are sent to relatively safe and benign environments to mitigate the risk of losing such an expensive piece of tax-funded equipment. Unfortunately, the most interesting places to explore on Mars are far from being safe and benign environments. Rocky outcrops, caves, deep craters, and dried river beds hold both high scientific interest and high risk. There are balancing priorities: risk the rover to perform valuable science; or risk the science return quality.

The solution is to send one or more comparatively cheap

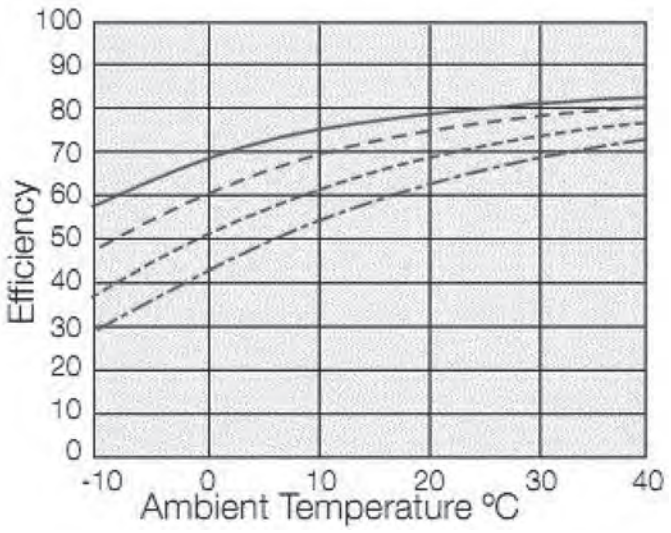


Fig. 6 Harmonic drive efficiency curves.
(Courtesy of Harmonic Drive LLC)

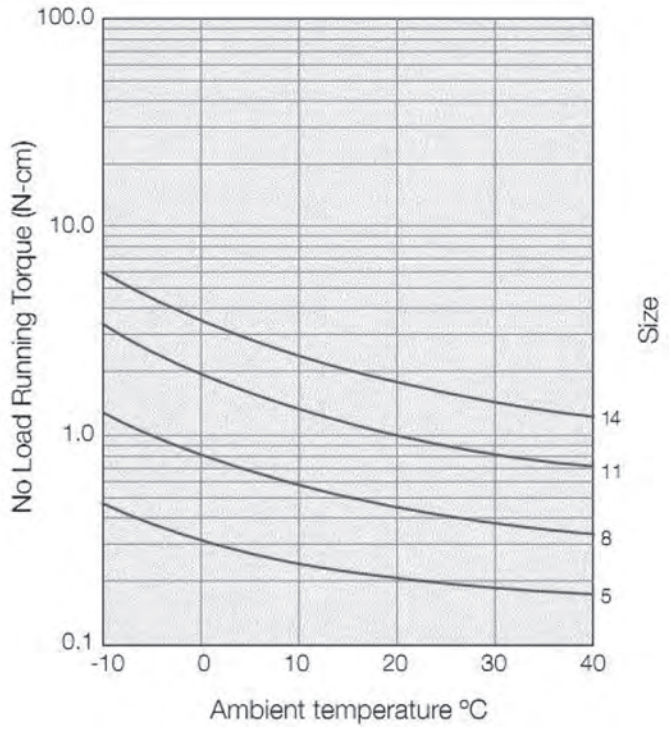


Fig. 7 Harmonic drive no-load curves.
(Courtesy of Harmonic Drive LLC)

and expendable micro-rovers to accompany the primary rover. These small scale rover scouts will be sent into the rocky outcrop or cave to take pictures and soil samples. They don't need to have the analysis equipment on-board; they just need to get the samples back to the primary rover. The scouts will also investigate deep craters. Should the scout be lost, it's unfortunate; however the main vehicle will continue to perform its mission. The remaining scouts, if any, will fill the void and explore the high risk areas and fetch the high scientific yield samples.

We propose sending one or more micro-rovers derived from *Kapvik* along with *ExoMars* to sample and fetch images, soil and atmospheric samples from the hostile environments on Mars. By qualifying off-the-shelf components, SEEG is lowering the costs for developing this class of rover. While

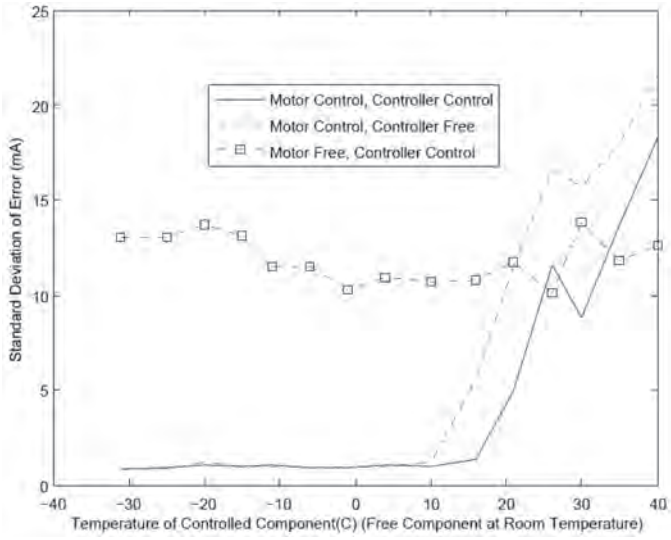


Fig. 8 Results of the Motor and Controller in Current Mode for Different Temperatures.

it is true that COTS components will not function in the full range of Martian environment conditions, the reduced cost associated is an acceptable trade-off for the scientific return gained.

6. EXPANDING THE TEMPERATURE RANGE OF COTS COMPONENTS

The motivation is to apply the "Surrey Approach" to use off the shelf components in environments for which they are not rated. A mission to Mars in the region of the *Viking* landing site would require a rover temperature range of -15ºC to -110ºC [11]. Comparing this range to those of the components in Table 1 shows the necessity to expand the temperature range of these components. The value of the EPOS controller is the absence of the interface development phase as mention previously. A combination of passive and active thermal control is required to survive in this temperature range. However, it was straightforward to experimentally expand the temperature range for confident operation. In addition, the motor controllers are located in the cab with the rest of the heat dissipating electronics.

Harmonic drives have been used in low temperature space applications since the Apollo 15 Lunar Rover Vehicle [9]. The ones selected for *Kapvik* have a limited range compared to the motors due its standard lubricating grease. The manufacturer of the harmonic drive can supply the units with grease rated for temperatures down to -30ºC thereby leading to improved performance.

7. CONCLUSIONS

The *Kapvik* rover is required to operate at -20ºC to meet the temperatures expected in the Canadian arctic and at equitorial Mars. The results of the temperature tests on Maxon motors and controllers show that the assemblies can operate below their published limits. The COTS components will even continue to operate at -30ºC. The limiting factor is the increasing viscosity of the lubrication in the gearing as the temperature decreases as experienced by the MFEX team. Changing the lubrication to a commercially available low temperature grease will improve the performance at -30ºC and perhaps even lower. In addition to satisfying the temperature requirements for the *Kapvik*

micro-rover, these results will allow Carleton's SEEG team to continue to develop mechanisms and hardware for Mars surface exploration technology.

ACKNOWLEDGEMENTS

The authors would like to thank the SEEG members (Tim Setterfield, Marc Gallant, Rob Hewitt, Brian Lynch, Adam Mack, Helia Sharif, Adam Brierly, Jesse Heimstra, and Cameron Frazier) for their tireless contributions to the development and testing of the *Kapvik* instrumented chassis and mobility system; without their hard work and dedication research performed for this paper would not be possible. *Kapvik* was funded by the Exploration Surface Mobility project of the Canadian Space Agency and administered by MPB Communications Inc. Ryerson University, the University of Toronto Institute for Aerospace Studies, MacDonanld Detwiller and Associates, and Xiphos Technologies Inc contributed to *Kapvik*'s development.

REFERENCES

1. D. Shirley and J. Matijevic, "Mars Pathinder Microrover", *Autonomous Robots*, **2**, pp.283-289, 1995.
2. S.W. Squyres, R.E. Arvidson, D. Bollen, J.F. Bell, J. Bruckner, N.A. Cabrol *et al.*, "Overview of the Opportunity Mars Exploration Rover Mission to Meridiani Planum: Eagle Crater to Purgatory Ripple", *Journal of Geophysical Research*, **111**, E12, 2006.
3. R.E. Arvidson, J.F. Bell, P. Bellutta, N.A. Cabrol, J.G. Catalano, J. Cohen *et al.*, "Spirit Mars Rover Mission: Overview and selected results from the northern Home Plate Winter Haven to the side of Scamander crater", *Journal of Geophysical Research*, **115**, E7, 2010.
4. B. Muirhead, "Mars rovers, past and future", in *2004 IEEE Aerospace Conference Proceedings*, IEEE, USA, 2004, pp.128-134.
5. T.P. Rivellini, "Mars rover mechanisms designed for Rocky 4", *The 27th Aerospace Mechanisms Symposium*, NASA Ames Research Center, Moffett Field, California, 12-14 May 1993.
6. J. Chottiner, "Simulation of a Six Wheel Martian Rover Called the Rocker Bogie", MSc Thesis. The Ohio State University; 1992.
7. K. Yoshida, T. Watanabe, N. Mizuno, G.Ishigami, "Slip, Traction Control, and Navigation of a Lunar Rover", *Proceedings of the 7th International Symposium on Artificial Intelligence, Robotics and Automation in Space*, May 2003.
8. G. Ishigami, A. Miwa, K. Nagatani and K. Yoshida, "Terramechanics based model for steering maneuver of planetary exploration rovers on loose soil", *Journal of Field Robotics*, **24**, pp.233-250, 2007.
9. K. Ueura and R. Slatter, "Development of the Harmonic Drive Gear For Space Applications", in *Proceedings of the 8th European Space Mechanisms and Tribology Symposium*, **Vol. 438**, p.259, 1999.
10. I. Schafer, P. Bourlier, F. Hantschack, E. Roberts, S. Lewis, D. Foster *et al.*, "Space Lubrication and Performance of Harmonic Drive Gears", in *Proceedings of the 11th European Space Mechanisms and Tribology Symposium*, ESA, 2005, pp.65-72.
11. J.F. Murphy, C.B. Leovy and J.E. Tillman, "Observations of Martian surface winds at the Viking Lander 1 Site", *Journal of Geophysical Research*, **95**, pp.14555-14576, 1990.

(Received 21 February 2012; Accepted 13 November 2012)

* * *

IN-ORBIT CONSTRUCTION WITH A HELICAL SEAM PIPE MILL

NEILL GILHOOLEY
16 Fleeman Grove, Nottingham, NG2 5BH, UK
Email: neillgilhooley@hotmail.co.uk

The challenges of building large structures in space, and in particular a torus habitat, require novel processes. One potential method is to manufacture helical seam (also called spiral) pipe in orbit using a pipe mill. These machines turn rolls of steel or alloy into fully formed, welded and inspected pipe, pressure vessels and silos of various diameters. Pipe mills are highly automated and efficient in a factory environment and are increasingly being used for in-situ repair. By constructing in-orbit (on-orbit assembly) the launch vehicle can supply full payloads of compact, robust rolls of material; and the installation design is less restricted by fairing constraints and modular limitations. The use of a pipe mill is discussed as a possible construction method, for comparison an example design envelope is shown and further pipe mill products are considered.

Keywords: In-orbit construction, pipe mill, habitat, torus, rotating wheel space station

1. INTRODUCTION

Significant decisions in space exploration to date have been taken over direct versus indirect modes [1], with the former proposing a single launch vehicle powerful enough for the task and the latter advocating some level of assembly in orbit from multiple launches.

The vast majority of installations in space have been lifted in single launches. With the space stations of the Salyut programme, Mir, and culminating in the International Space Station (ISS), far larger installations have been assembled in orbit using modules from multiple launches. However it is still the case that each individual module is limited to the capabilities of the launch vehicle, whilst adding the complexity of in-orbit assembly.

These long missions have allowed us to study the potentially debilitating effects on man of prolonged microgravity. High equipment failure rates have also been experienced from lack of heat convection [2]. In the decades preceding manned missions, many designs intended to mitigate these effects with a rotating wheel habitat to create artificial gravity by reaction to centripetal acceleration [3]. For an interplanetary mission, such as a voyage to Mars, it would be advantageous to provide artificial gravity for the health and wellbeing of the crew.

Without doubt a torus is more difficult to assemble than a monolithic structure, asking the question: how to sub-divide, strengthen and squeeze such an installation through the launch vehicle conduit from earth to orbit? Wernher von Braun sought to address this with several collapsible modules [4]. By the late 1950's his attitude to space station design (e.g. Fig. 1) was "Let's envision a space station and what [it] is made up of, what it can perform and not worry too much about how we would get it up there" [5]. Yet the construction method is key to the feasibility of such a space station, and this challenge requires the development of tools, subject to risk-benefit analysis, of absolute reliability and severe mass limitations. In this discourse an in-orbit construction method is considered to address both monolithic (tubular) and rotating wheel (toroidal) structures.

2. HELICAL SEAM PIPE MILLS

A very successful automatic manufacturing technique in terrestrial applications is the helical seam pipe mill, also called a spiral pipe mill. Two-thirds of steel tube production is by welded tube mills [6]. A tool of this efficiency, with high production standards and limited human intervention, is a strong candidate for in-orbit construction. Such a mill could manufacture tubular and toroidal structures from alloy and composites, with dimensions largely independent of the launch vehicle fairing size.

Helical seam pipe mills use rolled steel or alloy that is uncoiled, aligned and rolled by one internal, and a cage of external rollers to create a tube, welded internally and externally and then inspected (ultrasound and X-ray) as part of a continuous process. When one roll is finished another is welded on without interruption and sections may be automatically cut to length. The pipe may also be corrugated.

The pipe dimensions are produced according to:

$$\sin\alpha = \frac{W}{\pi d}$$

Where α is the skelp (strip) feed angle, W is the skelp width and d the tube diameter (minor axis for a torus) as per Fig. 2 [6].

Silos of large diameters, up to 25 m, can currently be manufactured in situ with lightweight equipment. There are also growing developments of construction and repair in demanding environments by helically wrapping pipelines; and internally installing liners underground, such as [7], a technique that might have applications in Martian lava tubes.

3. IN-ORBIT CONSTRUCTION

3.1 Tubular Products

The simplest form that the helical seam pipe mill can

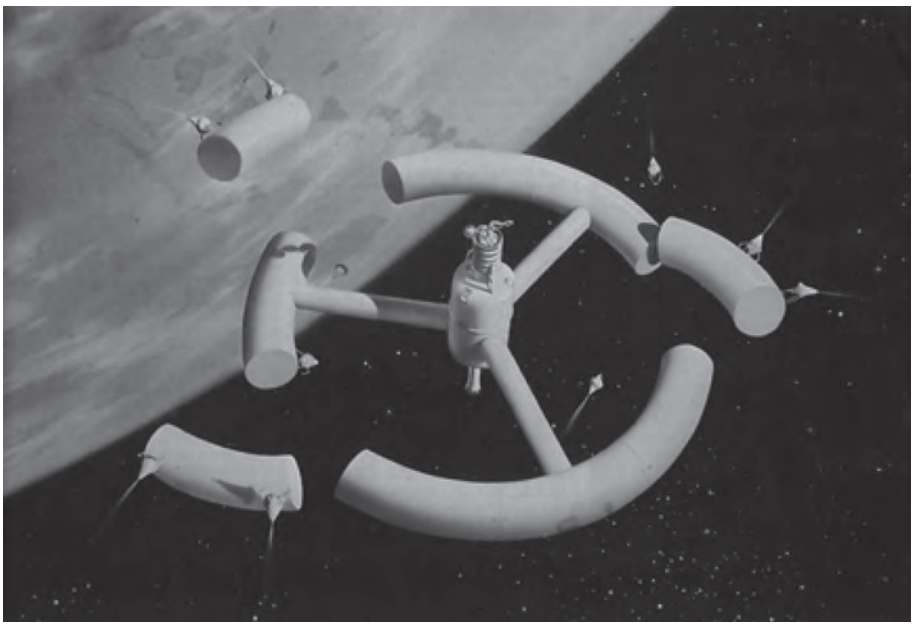


Fig. 1 von Braun space station.
(© Disney 1955)

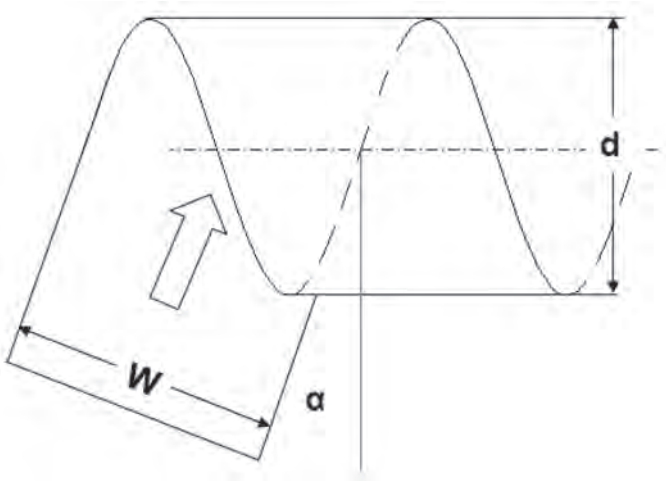


Fig. 2 Helical pipe.

manufacture is a tube. An in-orbit tubular build would start with the construction of the mill from modules lifted into orbit.

The mill consists of, say a hex frame larger than tube diameter *d*, with rollers to form helical pipe. This would be assembled by extra-vehicular activity (EVA). Specially prepared alloy and composite rolls, taken to orbit as dense payloads, are lined up to feed the mill. Each is tailored for the job with pre-prepared features and edges to minimise the task in-orbit.

First an inner skin of tubular alloy is rolled and internally and externally welded [8] to create an uninterrupted internal volume. A sensor head follows to inspect the integrity of the weld. This part of the process is largely automated. The mill can be positioned by rotation and translation to any point on the tube by its rollers, and could act like the Mobile Transporter aboard the ISS, carrying a robot arm.

Subsequent passes along the tube can be made adding additional layers, such as carbon fibre for strength. One preferred construction for Micro-Meteoroid Orbital Debris (MMOD) protection is the so-called 'stuffed Whipple shield' [9] with a blanket equidistant between two alloy skins. This blanket consists of a ceramic fabric layer (e.g. Nextel) backed by a high-strength fabric layer (e.g. Kevlar), and may also incorporate multi-layer insulation (MLI).

To separate inner and outer skins the mill could apply a helical strip stand-off normal to the skins, resembling a vortex shedding strake on a tall chimney stack. A consequence for the outer skin is that the internal roller would need to be narrower than the inter-strake gap. One possible method of suspending the blanket uses folds in the strake that collapse when compressed by rollers, crimping the blanket on both sides and affording a protective overlap.

By dint of the rollers being set to a larger diameter a second, outer skin can be rolled and welded externally. This seam can be out of phase with the inner for reasons of strength, or rolled with the opposite chirality.

The pipe mill would therefore make several passes over a multi-layer section during construction, the relative rotation dependent upon the changing mass distribution. After the automated processes, the provision of a rotating welding head also allows flanges, frames and supports to be added as required, under human guidance.

3.2 Torus Habitat

To create a torus habitat, sinusoidally edged skelp and Z-axis rollers would be used to produce an approximately curved tube, of major diameter *D*. Here the 3D problem is reduced to a single weld point on a continuous seam, rather than a series of discrete processes where the relative positions of all the elements must be co-ordinated. On completion of the structure the interior is loaded with equipment and the ends are joined. Composite decks may be used, or alloy with ribs pinched by the rollers.

Using an automated machine to construct a tested pressure vessel allows EVA activity to be kept to a minimum. Only after the habitat is pressurised would the crew be required to start on the next phase of the construction, namely fitting-out with the installation of equipment. This work would take place in a 'shirt-sleeve' environment, at either zero-g for the movement of heavy payload racks, or with artificial gravity for crew comfort.

JBIS

Journal of the British Interplanetary Society

JBIS SURVEY

In the last year many changes have been made to the Journal and we are also embarking on process improvement. It is now time to start thinking about making changes to the Journal to ensure it reflects what you the readers expect to see. So, we have constructed a short survey and we would be grateful if you could fill this in. We are very grateful for your assistance and continued loyalty to *JBIS*.

You can also complete this survey online at www.surveymonkey.com/s/Q5XTT7M

Please leave your name, phone number and email address if you are willing to be contacted by the Editorial team?

Name:..... Telephone Number: .....

Email:.....

1. How do you rate the quality of papers that appear in *JBIS*?

(a) Poor (b) Satisfactory (c) Average (d) Good (e) High

2. How do you think we are doing to ensure a diversity of papers that appear in *JBIS* covering the span of astronautics?

(a) Poor (b) Satisfactory (c) Average (d) Good (e) High

3. How do you rate the job we are doing in managing your Journal?

(a) Poor (b) Satisfactory (c) Average (d) Good (e) High

4. Please mark below an approximate percentage in the order of your preference for what phrase you feel best describes what the majority of *JBIS* papers should reflect. Please ensure your total adds up to 100%

(a) Speculative Conjecture (b) Visionary and Bold (c) Credible Concepts

(d) Realistic and Near-term (e) Application Proven

5. Are there subjects that you feel *JBIS* does not currently cover but which we should be?

6. Do you subscribe to any other astronautical, astronomy or space based Journals?

(a) Yes (b) No

7. If you answered yes to the above question, please name the Journal and describe how it compares to *JBIS*.

8. What do you like/dislike about *JBIS*?

9. Are you aware that you can now purchase papers from the *JBIS* web site (JBIS.org.uk)?

(a) Yes (b) No

10. Do you have any other comments or suggestions for improving *JBIS*?

Please return the survey to:
***JBIS* Survey, The British Interplanetary Society, 27/29 South Lambeth Road, London, SW8 1SZ**


THE BRITISH
1933
2013
INTERPLANETARY SOCIETY

The mill remains with the installation, the rotating sensor package continually inspecting the exterior of the torus and the mill proper on standby in case patch repairs are required.

3.3 Mixed Construction

In reality a mixed, hybrid construction method may be likely, utilising modular and mill methods with alloy and composites. For example two or three equispaced pre-fabricated ('tin can') modules could be used as part of the torus at spoke junctions. Such a module can be robustly built and may be fitted with windows, an EVA airlock, pressure doors etc. In the construction phase the mill can be attached to one end of such a module, reducing the weight of the mill.

The tube diameter could be large enough to be the entire installation (space station or spacecraft), a walkway of say 3 m in diameter, or merely a strength member, to which inflatable modules are attached. Similarly different tube diameter segments may be combined to provide large habitat volumes and a continuous corridor through inflatable greenhouses etc.

4. EXAMPLE DESIGN ENVELOPES

The straightforward relationship between area and volume for a torus is shown in Fig. 3 for various tube diameters *d* (minor axis). The designer may operate within the shaded design envelopes, limited by rate of rotation and tangential velocity for crew adaptability to artificial gravity [10], providing a free hand in terms of mass, tube and torus diameter, deck area and pressurised volume. However pre-fabricated structures are here limited to *d* 4.5 m by launch vehicle fairings, and inflatable designs [11] to an expanded diameter of 6.7 m.

Also shown in Fig. 3, as points, are three proposed space stations: Potočnik's of 1929 [3], von Braun's of 1952 and an octagon made from Shuttle External Tanks of the type described in reference [12] from the 1990's.

The mass of an example design could be reduced for the same volume by selecting a larger tube diameter *d*, putting us well into the mill design envelope. If this were as large a diameter as our chosen crew adaptability criteria permits, then there is a mass saving of about a third when hoop stress is taken into account. Alternatively, the example could be translated to the boundary for the same mass, again making an adjustment for hoop stress, increasing the internal volume by a half.

5. ADVANTAGES & APPLICATIONS

The advantages of the mill design over the modular (pre-fabricated or inflatable) designs are:

i) A larger volume is possible for a given mass, less the mass of the mill itself

ii) Installation components can exceed payload dimensions

iii) Large diameter shields can be added, bringing weight savings without impacting on the internal volume

iv) The installation can be designed for a low-g environment and will be lighter than modules that must withstand the rigours of launch

v) Fuller utilisation of each lift can be achieved closer to the average payload density of the launch vehicle. Using high density rolls 95% of a typical payload volume is free for a trade-off with lower density materials, such as carbon fibre decks and interiors

vi) The material can be conveniently divided into compact mass units. In particular Reusable Launch Vehicles (RLV), such as spaceplanes, promise to dramatically reduce the cost per kg to Low Earth Orbit (LEO), but the emphasis may be on payload mass with volume curtailed by hypersonic aerodynamics and structural considerations, or insist on g-hardened payloads

vii) A single automated process minimises EVA

Various tests can be made during construction, including weld inspections, pressure tests, activation of the monitoring system and EVA inspections. However a key deficiency with in-orbit construction is the limited test regime. Installation modules are among the most thoroughly tested structures made by man and this would be impossible to approach in orbit.

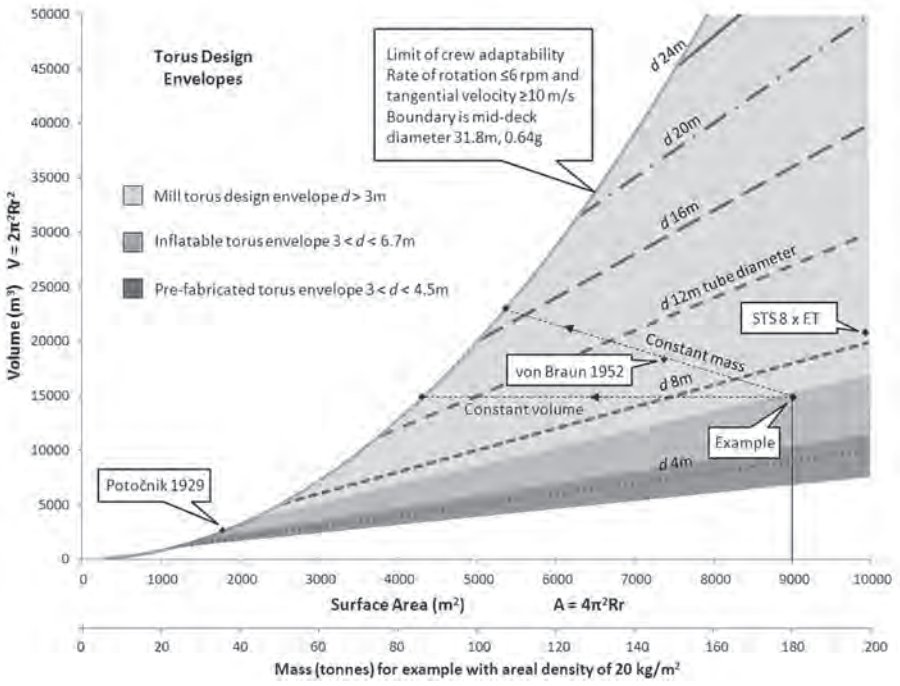


Fig. 3 Simplified Torus design envelopes.

A modular torus is less efficient because:

i) Multiple cylindrical modules result in a pseudo-torus of chords with varying centripetal accelerations

ii) Such a shape creates certain structural challenges over a continuous torus. Additional mass would be required to connect or flange each module together, without adding to the volume of the habitat

iii) Inflatable modules are typically necked to a core making for a lower volumetric efficiency, particularly if a curved interior is fitted

In short, the modular design envelopes and lifting plans are restrictive.

Almost any construction may benefit from tubular products at different diameters, such as:

i) Cylindrical housings or more complicated, non-circular cross-sections, and toroids

ii) 'Double-helix' sections with large, diamond-shaped, greenhouse windows

iii) Spars, spokes, frames and trusses including latticework for photovoltaic arrays

iv) Tanks for fluids and propellants

v) Pipework, heat pipes, fin tubes, heat exchangers; and plastic 'welded' conduits for internal applications

vi) Adding one or more large diameter Whipple shields to a pre-fabricated habitat. This is among the more pragmatic options, using a thoroughly tested module made to the limits of the launch vehicle, and adding large diameter shielding in orbit

vii) A large vessel could be made to enclose satellites for manned servicing, or even for asteroid capture

On the surface of the Moon or Mars the mill could manufacture many of the components of a base, such as:

viii) Habitats and interconnecting walkways buttressed with regolith

ix) 'Polytunnels' as greenhouses, possibly with a reinforcing wire in the seam

x) Well casings, pipelines, piles, silos and tanks for the storage of In Situ Resource Utilisation (ISRU) products

xi) Slender spars fitted with terminations or flattened and drilled in order to construct a geodesic framework for spheres, domes or parabolic reflectors

xii) A pressurised section could be produced as a land vehicle cab, fitted to a chassis

xiii) The mill could make the cylinder and blade of a continuous process cement mixer, and spend its last weeks turning this on its rollers to produce hundreds of tonnes of concrete (or lunarcrete, 'regocrete' etc)

These structures can be delivered 'flat-packed' as compact and robust alloy and composite rolls to the surface of the planet.

6. CONCLUSIONS

Central to mankind's progress has been the development of tools that have allowed us to prosper in diverse environments. Construction tools are required for large installations, including those that may be suitable for prolonged manned missions such as voyages to Mars.

Ideally an installation design would not be limited by fairing dimensions, nor by the stresses of launch; but be constructed to an uncompromised design for a microgravity environment. In addition a rotating wheel installation (space station or spacecraft) is an attractive option to create artificial gravity. However, such a structure is difficult to construct from modules. Automated pipe mills work well on Earth but an in-orbit helical seam pipe mill would have to deliver a faultless product.

A wide variety of tubular products could be manufactured by an effective construction robot like a pipe mill, enabling large design envelopes. Chief amongst these products, efficient toroids can be constructed to make large volume habitats with artificial gravity.

ACKNOWLEDGEMENTS

The author is grateful to Disney Enterprises, Inc, for permission to use the image in Fig. 1, and to *JBIS* for their assistance.

REFERENCES

1. C.G. Brooks, J.M. Grimwood and L.S. Swenson Jnr, "Chariots for Apollo", NASA SP-4205, Ch.3-2, 1979.
2. K. Legner, "*Humans in Space & Space Biology*", United Nations Office for Outer Space Affairs, pp.81-83, 2004.
3. H. Potočnik (H. Noordung), "*The Problem of Space Travel*", 1929, translated NASA SP-4026, p.102, 1995.
4. "Giant Doughnut Proposed as Space Station", *Popular Science Monthly*, p.120, October 1951.
5. A.J. Dunar and S.P. Waring, "*Power to Explore: A History of Marshall Space Flight Center 1960-1990*", pp.529-530, 1999.
6. K-H. Drensing and B. Sommer, "*Steel Tube and Pipe Manufacturing Processes*", pp.55-62, 2004.
7. K. Hanks *et al*, "*City of Los Angeles Conducts a Test Installation of a Non-Circular Wound Pipeline Rehabilitation Method*", North American Society for Trenchless Technology, Paper B-3-04-1, 2005.
8. T.W. Eagar *et al*, "In-Space Welding, Visions and Realities", *Thirtieth Space Congress "Yesterday's Vision is Tomorrow's Reality"*, Cocoa Beach, Florida, USA, 27-30 April 1993.
9. E.L. Christiansen, "*Handbook for Designing MMOD Protection*", NASA/TM-2009-214785, 2009.
10. T.W. Hall, "Artificial Gravity and the Architecture of Orbital Habitats", *JBIS*, **52**, pp.290-300, 1999.
11. B. Rubanovich, "Bigelow Slideshow", International Space Development Conference, 2011.
12. T.C. Taylor, "Centrifugal Gravity Habitation Torus Constructed of Salvaged Orbital Debris", United States Patent US 6,206,328, 1998

(Received 23 February 2012; Revision 21 November 2012 & 10 May 2013; Accepted 28 May 2013)

JBIS, Vol. 66, pp.171-177, 2013

THE EFFECT OF PROBE DYNAMICS ON GALACTIC EXPLORATION TIMESCALES

DUNCAN H. FORGAN*, SEMELI PAPADOGIANNAKIS AND THOMAS KITCHING
Scottish Universities Physics Alliance (SUPA), Institute for Astronomy, University of Edinburgh, Blackford Hill, Edinburgh, EH9 3HJ, UK.
Email: dhf@roe.ac.uk*

The travel time required for one civilisation to explore the Milky Way using probes is a crucial component of Fermi's Paradox. Previous attempts to estimate this travel time have assumed that the probe's motion is simple, moving at a constant maximum velocity, with powered flight producing the necessary change in velocity required at each star to complete its chosen trajectory. This approach ignores lessons learned from interplanetary exploration, where orbital slingshot maneouvres can provide significant velocity boosts at little to no fuel cost. It is plausible that any attempt to explore the Galaxy would utilise such economising techniques, despite there being an upper limit to these velocity boosts, related to the escape velocity of the object being used to provide the slingshot. In order to investigate the effects of these techniques, we present multiple realisations of single probes exploring a small patch of the Milky Way. We investigate 3 separate scenarios, studying the slingshot effect on trajectories defined by simple heuristics. These scenarios are: i) standard powered flight to the nearest unvisited star without using slingshot techniques; ii) flight to the nearest unvisited star using slingshot techniques, and iii) flight to the next unvisited star which provides the maximal velocity boost under a slingshot trajectory. We find that adding slingshot velocity boosts can decrease the travel time by up to two orders of magnitude over simple powered flight. In the third case, selecting a route which maximises velocity boosts also reduces the travel time relative to powered flight, but by a much reduced factor. From these simulations, we suggest that adding realistic probe trajectories tends to strengthen Fermi's Paradox.

Keywords: Fermi Paradox, SETI, interstellar exploration, probe dynamics

1. INTRODUCTION

Fermi's Paradox remains an important cornerstone of modern thinking on extraterrestrial intelligence (ETI). It taxes most, if not all attempts to formulate an optimistic perspective on the frequency of alien civilisations in the Galaxy both in space and time.

Detailed reviews of the Paradox can be found in [1, 2 and 3]. The Paradox rests on the current absence of ETI in the Solar System (what [4] refers to as "Fact A"). This absence runs counter to estimated timescales for intelligent species to colonise of the Galaxy - what [3] (and references within) refers to as the Fermi-Hart timescale (see e.g. [4, 5]):

$$t_{FH} = 10^6 - 10^8 \text{ yr} \quad (1)$$

This is compounded by the fact that the age of the Earth is at least an order of magnitude higher, and the median age of terrestrial exoplanets is estimated to be a further 1 Gyr older than the Earth [6]. It appears the inexorable conclusion one must draw is that ETIs do not exist, otherwise we would have detected their presence in the Solar System.

It is common for attempts to resolve the Paradox to speculate on the motivation or sociological make-up of ETIs - for example, one solution suggests that the Earth has attained a protected status and must not be disturbed, often known as the Zoo Hypothesis (e.g. [7]). While flaws can be exposed in these types of hypothesis (see e.g. [8]), there are many solutions that are simply unfalsifiable, and while they cannot be ruled out, they cannot be currently considered as scientifically meritorious. Until conclusive data is compiled on something as esoteric as extraterrestrial sociology, it is more worthwhile to focus on potential physical constraints for extraterrestrial contact.

Weaker formulations of the Paradox merely rest on ETIs practising interstellar communication rather than interstellar colonisation (e.g. [9]) - stronger formulations use self-replicating Von Neumann probes to explore the galaxy at an exponential rate. It is this process of self-replication (or colonies carrying out subsequent autonomous colonisation) that allows t_{FH} to be small enough for the Paradox to be robust. While there have been many arguments for and against the use of self-replicating probes (e.g. [5, 10, 11, 12]), we wish to focus instead on a more fundamental aspect of probe exploration that has not been addressed fully.

The Paradox leans heavily on the dynamics of interstellar flight, and the motivations of the extraterrestrial intelligences (or ETIs) that drive the exploration. Sagan [13] expounds a framework for a cadre of civilisations visiting star systems using relativistic interstellar flight. Under these assumptions, the visiting rate for main sequence stars could be as high as once per ten thousand years (although it makes relatively optimistic assumptions about the number of civilisations forming the cadre). The associated problems of population growth and carrying capacity are also important drivers for continual exploration, as was explored by Newman *et al.* [14] using nonlinear diffusion equations. The stipulation that ETIs practise zero population growth can alter t_{FH} by several orders of magnitude.

At a more fundamental level, Bjoerk [15] investigated probe exploration in a schematic model of the Galaxy, with an exponentially declining stellar surface density distribution. Each host probe visits a subset of the Galactic stellar population, containing 40,000 stars. The host then releases 8 sub-probes, which explore this subset, and then dock with the host before travelling to a new Galactic sector. Each probe has a constant velocity of 0.1c, and travels to its nearest unvisited star under powered flight. There is no attempt to optimise the trajectory of the sub-probes beyond this simple heuristic. Optimising its trajectory is an instance of the "travelling salesman" problem [16, 17], an NP-hard problem which is computationally prohibitive to solve for large node numbers, as is the case in interstellar exploration.

Cotta *et al*. [18] extended this work by using algorithms to improve the trajectory of the probes, reducing the tour length by around 10%. They subsequently carried out a probabilistic analysis assuming many ETIs were carrying out exploration in this fashion. This allowed them to estimate the maximum number of ETIs exploring the Galaxy (N_{ETI}) that could do so while ensuring contact with Earth remained statistically unlikely. In the parameter space explored, they estimate that the upper limit is approximately $N_{ETI} < 10^{-3}$.

What is common to all these simulations of probe exploration is that they ignore or neglect lessons learned from Mankind's unmanned exploration of the Solar System. Simple powered flight is an inefficient means of travel, especially for probes using chemical or nuclear propulsion methods. Using slingshots inside gravitational potential wells allows the probe to produce relatively large Δv and alter its trajectory without expending fuel, and potentially boosting its speed relative to the rest frame of its starting position.

This has been utilised successfully by Mankind during its exploration of the Solar System, both in the ecliptic plane through the triumphal tour of the Voyager probes (see e.g. [19]), and even out of the ecliptic during the Ulysses mission [20].

In principle, this behaviour can be scaled up to the Galactic level, where the potential well of stars can now be used to provide Δv and increase the probe speed relative to the Galactic rest frame [21]. While this requires probes to be able to navigate extremely accurately, this would not appear to be an insurmountable obstacle provided the probe possessed sufficient autonomous computing power, and retained enough fuel for judicious course corrections.

In this paper, we investigate the effect of individual probe dynamics on the visitation timescale of a population of stars. We run a series of Monte Carlo Realisations (MCRs), where in each realisation a single probe traverses a path through a population of stars. The stellar population has a number density and velocity field akin to that of the Solar Neighbourhood. We run three separate sets of realisations, focusing on three basic scenarios:

1. A single probe, visiting stars under powered flight, with each leg of the route determined by finding the nearest unvisited neighbour (which we label **powered**),
2. As 1, except utilising slingshot trajectories to boost the velocity of the probe (which we label **slingshot**),
3. A single probe which selects the next star to travel to such that the velocity boost derived from a slingshot is maximal, which depends on the destination star's velocity relative to the current star (which we label **maxspeed**).

In section 2, we describe the mechanics of the slingshot maneouvres employed in this paper, as well as the setup of the simulations; section 3 describes the results of the three scenarios, and in sections 4 and 5 we discuss and conclude the work.

2. METHOD

2.1 The Dynamics of Slingshot Trajectories

A slingshot trajectory uses the momentum of the star it passes to gain or lose velocity, depending on the incident angle of the probe's approach. The probe therefore does not need to expend additional energy that would otherwise be required to complete a similar trajectory under powered flight. We briefly describe the mathematics of slingshot trajectories here: for more detail the reader is referred to [19] (in particular to Chapter XIII, section 4, equations 29 to 31).

The left panel of Fig. 1 describes one stage of the probe's trajectory when slingshots are used. The probe begins at Star 0, accelerating to velocity u_i (measured in the frame where Star 1 is at rest). For the probe to proceed to star 2 after the slingshot, its velocity must assume the vector u_f, again in Star 1's frame. The angle between u_i and u_f is δ. In this manoeuvre, the probe follows a hyperbolic trajectory, and

$$|u_i| = |u_f| \tag{2}$$

However, we are interested in the probe's velocity in the "lab" frame, i.e. a frame of reference where the Galactic Centre is at rest. The right panel of Fig. 1 shows the vector diagram connecting the initial and final velocities in both frames. As a result of adding the star's motion to u_i and u_f to create the Galactic frame velocities v_i and v_f, we can see that there is indeed an increase in the probe's speed, Δv, which we can deduce simply:

$$\Delta v = 2|u_i|\sin\left(\frac{\delta}{2}\right) \tag{3}$$

The change in the probe's momentum in this frame is balanced by a change in the star's momentum relative to the Galactic Centre. The fractional decrease in momentum the star experiences is so small that we can effectively regard it as negligible. The angle δ is related to the inward velocity as follows:

$$\tan\left(\frac{\delta}{2}\right) = \frac{GM_*}{r_c u_i} \tag{4}$$

where r_c is the distance of closest approach to the star. This places an upper limit on the value of δ and consequently Δv (see Discussion).

We assume the probe has a maximum velocity it can attain under its own power. This defines the probe's initial velocity as it travels from Star 0 to Star 1. Henceforth, it can travel with increasing speed as it undergoes a series of slingshot maneouvres. The magnitude of the boost achieved by the slingshot maneouvre is increased if the star's own velocity runs parallel to the probe's trajectory. Therefore, it is possible that a

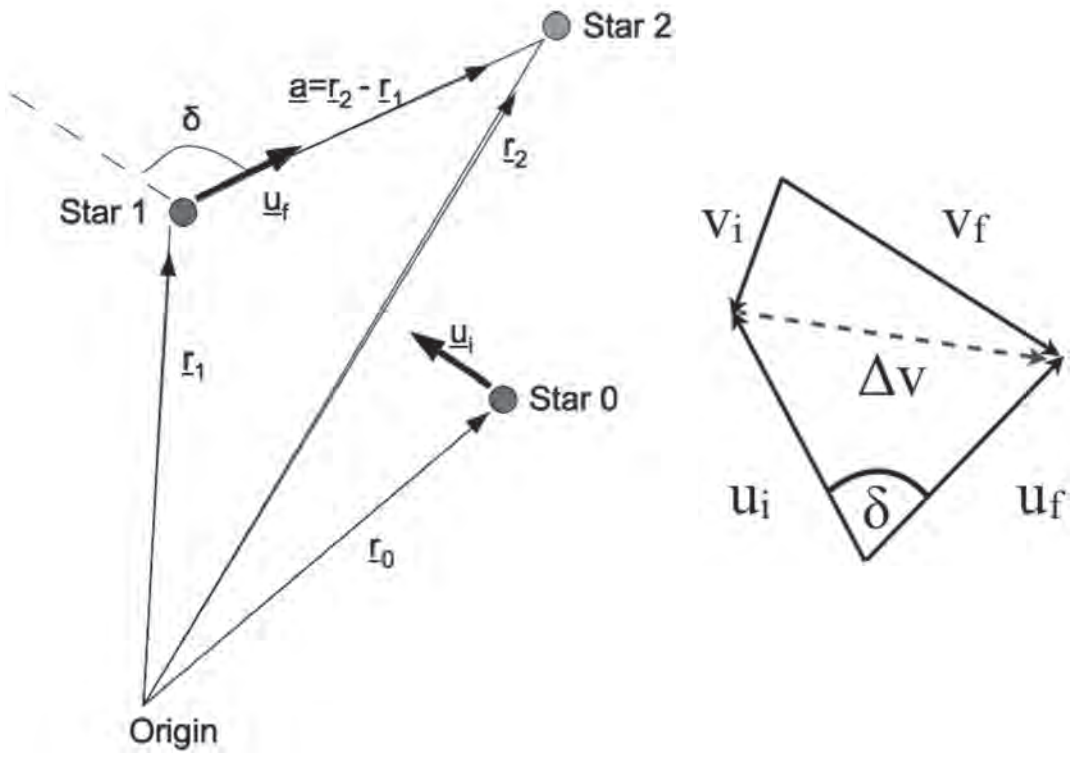


Fig. 1 The vector diagram above shows the path of the probe from the star's reference frame, changing its direction by an angle δ but not the magnitude of its velocity.

probe could choose a course based on the spatial velocities of stars relative to each other, such that it carries out slingshots with maximal Δv.

2.2 Simulating Probe Motion

We carry out three simulation scenarios, all in a small patch of the Galaxy containing one million stars, at a uniform density of 1 star per cubic parsec. The stars were set in a shearing sheet configuration to emulate the rotation curve of the Milky Way. For convenience, the stars were given velocity vectors for the slingshot calculation, but fixed in position. For each scenario, 30 realisations were carried out - this number represents a balance between reducing computational expense and maintaining a sufficiently small standard error arising from random uncertainties. As we will see in later sections, 30 realisations is sufficient to produce total travel times at accuracies of a few percent. In each realisation, the probe is allocated a different starting star.

The three scenarios are:

1. Powered flight to the nearest neighbour (**powered**). The probe travels from the starting star to the closest neighbour at its maximum powered velocity. The Δv is therefore fixed by the repeated deceleration and acceleration the probe makes at every stage of the trip.
2. Slingshot assisted flight to the nearest neighbour (**slingshot**). Here the path is identical to that taken when using powered flight due to the same method of choosing the next star. However, the probe need only expend Δv to accelerate to maximum velocity once, and does not need to decelerate, instead using the slingshot manoeuvre to repeatedly boost its maximum velocity. We assume that the δv expended by the probe to make course corrections to adopt a slingshot trajectory is negligible.
3. Slingshot assisted flight seeking the maximum velocity boost (**maxspeed**). This selects a different path entirely to the other two scenarios, seeking instead the course such that the relative velocity vector between the current and destination stars is large and negative, i.e. the destination star is moving toward the current star. This will give a larger velocity boost, but will in general require a longer path length to achieve it.

We select a relatively low maximum velocity of 3×10^{-5} c, where c is the speed of light in vacuo. This is comparable to the maximum velocities obtained by unmanned terrestrial probes such as the Voyager probes [22]. Admittedly, the Voyager probes achieved these speeds thanks to slingshot trajectories, so the top speed of human technology under purely powered flight is unclear. To some extent, the maximum powered speed of the probe is less important - increasing or decreasing this maximum will simply affect the absolute values of the resulting travel times in a similar fashion. What is more important is the relative effect of changing the propulsion method and/or the trajectory.

3. RESULTS

In Table 1 we summarise the three scenarios in terms of the total travel time to traverse all stars in the simulation, averaged over all 30 realisations. As the probe velocity we have selected is particularly low, the total travel times are quite large. In fact, two scenarios (powered and maxspeed) have travel times longer than the current age of the Universe (1.37×10^{10} yr), despite the standard error on the mean (which is generally a few percent of the mean). Therefore, it is unlikely that probes travelling at maximum velocities for current human technology can explore even a small fraction of their local neighbourhood in a timely fashion. We will investigate the reasons for this in more detail in the following sections.

TABLE 1: *A Summary of the Three Scenarios Investigated in this Work.*

Simulation	Average Total Travel Time (yr)	Standard Error (yr)
Powered	4.51×10^{10}	2.82×10^{7}
Slingshot	3.99×10^{8}	3.51×10^{5}
Maxspeed	1.99×10^{10}	1.23×10^{7}

On the other hand, the slingshot scenario has a significantly shorter travel time than the others. Of course, this travel time is only for 10^6 stars, and the Galaxy itself contains around 10^{11} stars, so it still remains unlikely that probes of this type could explore the entire Galaxy before Mankind began to construct devices that could detect them. For the Fermi Paradox to hold, the initial probe velocity would need to be much larger, a fact that has long been obvious [13].

3.1 Powered Flight Only, Nearest Neighbour (Powered)

Figure 2 shows the travel time between each star in the powered case. Here we average over 30 realisations, with the mean drawn in black and the standard error on the mean plotted in grey. Note that as the density of stars is roughly constant throughout the box, and the Δv achievable by the probe is fixed by its own engines, the travel time between nearest neighbours is also reasonably constant. As the number of unvisited stars drops below a hundred thousand or so, the probe must travel longer distances to find an unvisited star, and hence the travel time increases towards the end of the simulation.

3.2 Slingshot Trajectory, Nearest Neighbour (Slingshot)

If we now allow the probe to make slingshot maneouvres to its nearest neighbour (slingshot), then we can see that the probe's behaviour changes significantly (Fig. 3). It follows the same course as before, but it can now boost its velocity at every stage of the journey (right panel), significantly reducing its travel time (left panel). As a result, it has increased its velocity by almost a factor of 100 throughout the simulation (a fact borne out by its total travel time being approximately 100 times smaller than the powered case). Even towards the end of its path, where the unvisited nearby stars reside at larger distances, the increased speed as a result of the slingshots allow the probe to cover these larger distances in a short time span. Note the large standard error on the mean - Δv is now a sensitive function of the probe's path (more specifically, the angle δ in Fig. 1). As the probe begins from a different starting star in each realisation, the probe's path is significantly different, resulting in a large spread of Δv at each stage of the journey. The probe's maximum velocity is eventually limited by the minimum value δ can take (see Discussion).

3.3 Slingshot Trajectory, Highest Δv (Maxspeed)

With the probe now attempting to maximise its Δv, it pursues a fundamentally different trajectory to the other two cases. The selection criterion for the next star of the journey is purely to maximise Δv - the distance to the next star is not considered. The left hand panel of Fig. 4 demonstrates the consequences of this. The travel time indicates protracted journeys between stars in the early stages of the probe's trip.

The spikes in travel time are of high significance, and represent the extent of the simulation box. No periodic boundary conditions are applied - as a result, the probe commonly selects destinations that are at the opposite end of the box, and as a result must traverse the box's entire length frequently. The amplitude of these peaks decreases as the probe begins to boost its speed strongly (right hand panel of Fig. 4). By selecting for maximal Δv, the probe can achieve velocity increases 2.5 times larger than the slingshot case.

4. DISCUSSION

4.1 Limitations of the Analysis

Before discussing the implications of these results, we should

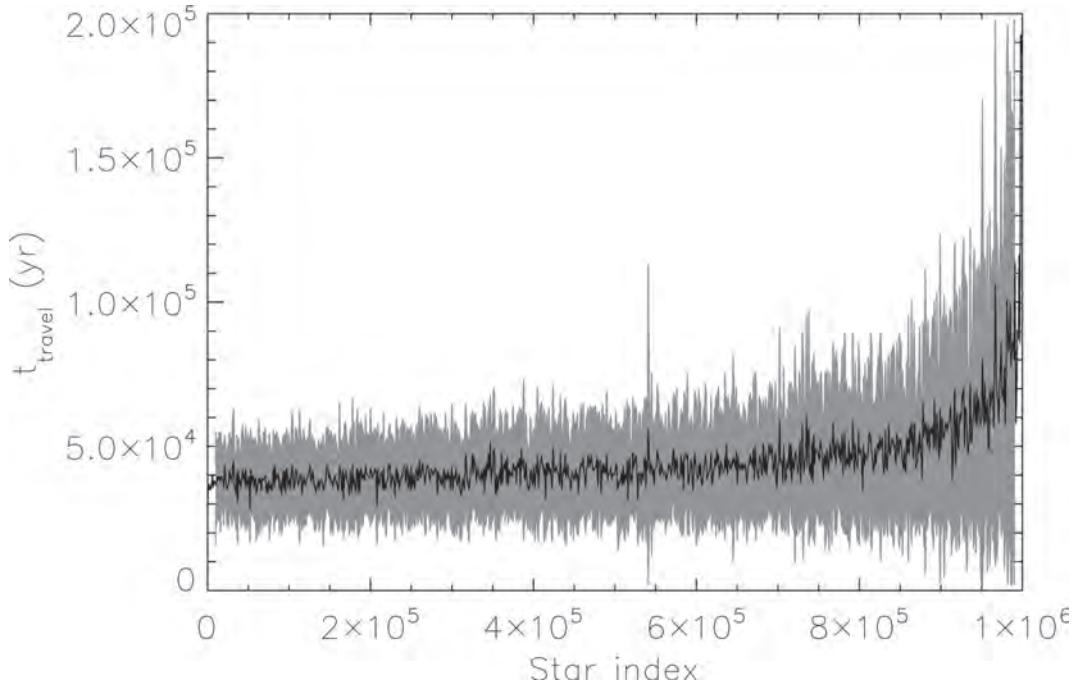


Fig. 2 The mean travel time between each star for probes in the powered simulations, averaged over 30 realisations. The grey shaded area represents the standard error on the mean.

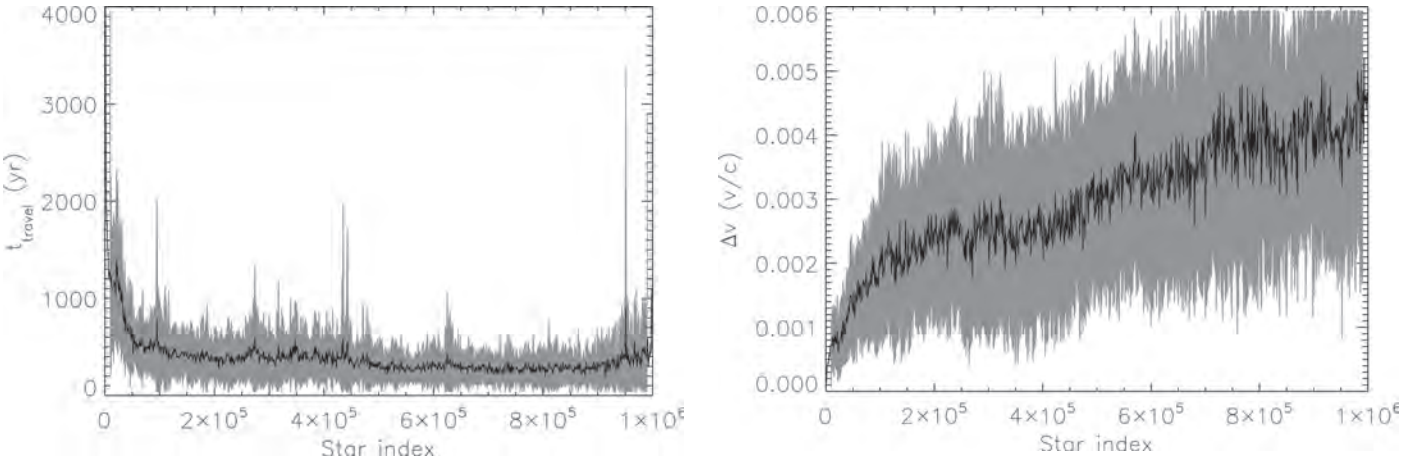


Fig. 3 Left: The mean travel time between each star for probes in the slingshot simulations, averaged over 30 realisations. The grey shaded area represents the standard error on the mean. Right: The mean Δv generated as a result of each slingshot maneouvre, with standard error on the mean shaded grey.

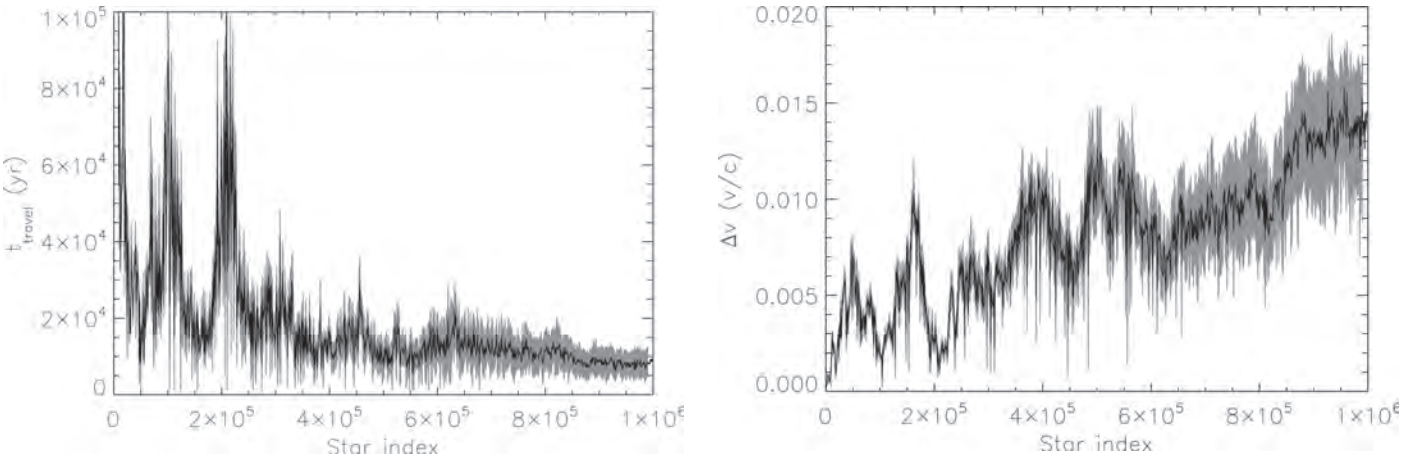


Fig. 4 Left: The mean travel time between each star for probes in the maxspeed simulations, averaged over 30 realisations. The grey shaded area represents the standard error on the mean. Right: The mean Δv generated as a result of each slingshot maneouvre, with standard error on the mean shaded grey.

first note some simplifications made to expedite the analysis.

Probably the most important simplification made was to fix the stars in position. While all stars possessed a velocity vector, the position vectors were never updated. As the aim was to create a large number of realisations at modest computational expense, it was felt that this assumption was reasonable. However, it does force us to jettison important dynamical aspects. For example, in the maxspeed case, the probe will select destination stars with relative velocity vectors which are large and negative, i.e the destination star and the current star are moving towards each other. This will reduce the travel time (just as positive relative velocity will increase the travel time). Future work in this area should investigate this effect.

The probe's motion was also simplified in several ways. The maximum Δv achievable by the probe is limited by the distance of closest approach that a probe can make to the star. From equation (4), for a star with mass M_* and an effective radial boundary R_{eff} that the probe cannot cross, there is a maximum value of rotation angle δ dependent on the inward stellarcentric velocity, u_i [19]:

$$\tan\left(\frac{\delta}{2}\right) = \frac{2GM_*}{R_{\text{eff}}\, u_i} \tag{5}$$

Substituting this into equation 3, we obtain:

$$\Delta v_{\max} = \frac{u_{esc}^2}{\frac{u_{esc}^2}{2u_i} + u_i} \tag{6}$$

where u_{esc} is the escape velocity at the radial boundary of the star:

$$u_{esc} = \sqrt{\frac{2GM_*}{R_{\text{eff}}}} \tag{7}$$

The maximum Δv achievable becomes quite negligible as u_i exceeds ~0.01c - for example, a star of one solar mass $1M_{odot}$ and $R_{eff} = R_\odot$ gives a maximum Δv of ~10^{-11} c. To improve this value, the probe would have to risk very close approaches with massive, compact objects such as neutron stars and black holes [23], which could present hazardous tidal forces upon the craft's hull.

We have ignored relativistic effects in this analysis. The probe achieves velocities of ~0.01c, which gives a Lorentz factor indicating that classical physics is an acceptable

approximation for this work. However, future studies that use a higher initial velocity should be cognisant of this.

$$\gamma = \frac{1}{\sqrt{1-\left(\frac{v}{c}\right)^2}} = 1.00005 \quad (8)$$

4.2 Implications for Fermi's Paradox

Let us now investigate possible constraints on Fermi's Paradox, and in particular, the Fermi-Hart timescale t_{FH}. If we assume a constant number density of stars, then we can estimate $t_{\rm FH}$ from the average total travel times calculated in this study:

$$t_{FH} \sim \left(\frac{3\times10^{-5}}{\frac{v_{\rm initial}}{c}}\right)\left(\frac{N_{star}}{10^6}\right)t_{\rm travel} \quad (9)$$

Assuming that the number of stars in the Galaxy, $N_{\rm star} = 10^{11}$, then at $v_{initial} = 3 \times 10^{-5}\ c$:

$$t_{FH} \sim 10^5 t_{\rm travel} \sim 10^{13} - 10^{15}\ {\rm yr} \quad (10)$$

This confirms that given the low initial velocity we selected, motivated by current speed records of human-manufactured probes, we are currently incapable of exploring the Galaxy inside the Hubble time with a single probe. However, we should still note that we are much stricter than we might need to be when imposing the maximum powered velocity of a spacecraft, and the relative decrease in travel time depending on probe dynamics is significant (as shown in Table 1). As is well known, if higher initial velocities are indeed possible, then the Fermi-Hart timescale becomes amenable. We can estimate the minimum velocity by rearranging equation (9) for $v_{initial}/c$:

$$\frac{v_{\rm initial}}{\rm c} \sim 3\times10^{-5}\left(\frac{t_{\rm travel}}{t_{FH}}\right)\left(\frac{N_{star}}{10^6}\right) = 3\left(\frac{t_{\rm travel}}{t_{FH}}\right) \quad (11)$$

i.e. even if $v_{initial} \sim c$, then

$$t_{FH} \sim 3t_{\rm travel} \sim 10^9 - 10^{11}\ {\rm yr} \quad (12)$$

which is still quite high (except for possibly the slingshot case, and even then the velocity boosts possible will be much smaller than those achieved at low speeds). From this, we appear to confirm previous calculations that even when probe dynamics are considered in more detail, one of two conditions must be satisfied for t_{FH} to be sufficiently low:

1. Faster than light travel is possible, or
2. Multiple probes are required [5, 12].

While option 1 has been explored theoretically and found to be possible for massless particles and for civilisations that can correctly manipulate space-time (see e.g. [24] for a review of these concepts), it requires the existence of exotic matter and energy sources, not to mention staggering technological prowess we do not yet possess. Option 2 would therefore appear to be the most plausible choice. Indeed, probes which carry out a series of stellar flybys without leaving significant evidence of their passage would not affect Fermi's Paradox at all, as individual star systems would only be able to detect these probes briefly (if at all). It would seem reasonable then to require that probes leave some sort of beacon or artifact in their wake as they pass through a system, to signal their existence and to strengthen Fermi's Paradox. The manufacture (or replication) of many beacons becomes an industrial problem on a scale similar to that of manufacturing multiple probes.

Given our current ability to manufacture large numbers of similar sized craft for terrestrial uses, it is not unlikely that we can adopt a similar approach to building probes. A simple calculation shows that producing 10^{11} Voyager-esque probes would allow humankind to explore the Galaxy in 10^9 years. Given that around 5×10^7 automobiles are produced each year globally [25], it seems reasonable to expect a coordinated global effort could produce the requisite probes within a few thousand years. If the probes are made to be self- replicating, using materials en route to synthesise copies, the exponential nature of this process cuts down exploration time dramatically [12].

Whether ETIs create a static fleet of probes launched from one source, or an exponentially growing fleet of probes replicated in transit from raw material orbiting destination stars, we argue that these measures strengthen the Fermi Paradox further when slingshot dynamics are included, but the relative strength of dynamics versus numbers are currently unclear. To answer this, we are repeating the analysis made in this paper with self-replicating probes to investigate whether optimised slingshot dynamics are even worthwhile in a cost-benefit scenario (Nicholson *et al*, in prep).

Having said this, probes carrying out slingshot trajectories in a larger, more realistic domain will be able to produce more realistic trajectories. This is another important avenue for further investigation.

5. CONCLUSIONS

We carried out Monte Carlo Realisation (MCR) simulations of a single probe traversing a section of the Galaxy, exploring the effect of slingshot dynamics on the total travel time. Three scenarios were explored: the standard scenario where probes travel under powered flight to their nearest neighbour (powered); traversing the same path, but utilising slingshot trajectories to boost the probe's velocity (slingshot); and a third scenario where the probe selects its next destination in order to maximise the probe's velocity boost (maxspeed).

We find that allowing the probe to make slingshots can reduce the probe's total travel time by two orders of magnitude. The velocity boost is typically additive, and as a result this velocity boost could presumably be marginally increased by allowing the probe to explore a larger domain and a higher quantity of stars.

The speed of the probe was selected to represent current human ability in unmanned spaceflight. The exploration time of probes moving at this speed is sufficiently large that humans would struggle to explore the Galaxy inside one Hubble time without mass production of probes. Our work shows that even when optimising probe trajectories are taken into account, the travel time remains quite large.

ACKNOWLEDGEMENTS

Duncan gratefully acknowledges support from STFC grant ST/J001422/1. Semeli gratefully acknowledges support from a Robert Cormack Bequest Vacation Scholarship (provided by the Royal Society of Edinburgh).

REFERENCES

1. G.D. Brin, "The Great Silence - the Controversy Concerning Extraterrestrial Intelligent Life", *QJRAS*, **24**, pp.283-309, 1983.
2. S. Webb, "*If the Universe is Teeming with Aliens– where is Everybody?: Fifty Solutions to the Fermi Paradox and the Problem of Extraterrestrial Life*", Springer, 2002.
3. M.M. Cirkovic, "Fermi's paradox: The last challenge for copernicanism?", *Serbian Astronomical Journal*, **178**, pp.1-20, 2009.
4. M.H. Hart, "Explanation for the Absence of Extraterrestrials on Earth", *QJRAS*, **16**, pp.128-135, 1975.
5. F.J. Tipler, "Extraterrestrial intelligent beings do not exist", *QJRAS*, **21**, pp.267-281, 1980.
6. C. Lineweaver, "An Estimate of the Age Distribution of Terrestrial Planets in the Universe: Quantifying Metallicity as a Selection Effect", *Icarus*, **151**, pp.307-313, 2001.
7. J. Ball, "The zoo hypothesis", *Icarus*, **19**, pp.347-349, 1973.
8. D.H. Forgan, "Spatio-temporal constraints on the zoo hypothesis, and the breakdown of total hegemony", *International Journal of Astrobiology*, **10**, pp.341-347, 2011.
9. L.K. Scheffer, "Machine Intelligence, the Cost of Interstellar Travel and Fermi's Paradox", *QJRAS*, **35**, pp.157-175, 1994.
10. C. Sagan and W.I. Newman, "The Solipsist Approach to Extraterrestrial Intelligence", *QJRAS*, **24**, p.113, 1983.
11. C.F. Chyba and K.P. Hand, "ASTROBIOLOGY: The Study of the Living Universe", *Annual Review of Astronomy and Astrophysics*, **43**, pp.31-74, 2005.
12. K.B. Wiley, "The Fermi Paradox, Self-Replicating Probes, and the Interstellar Transportation Bandwidth", *arXiv e-prints 1111.6131*, Nov. 2011.
13. C. Sagan, "Direct contact among galactic civilizations by relativistic interstellar spaceflight", *Planetary and Space Science*, **11**, pp.485-498, 1963.
14. W.I. Newman and C. Sagan, "Galactic civilizations: Population dynamics and interstellar diffusion", *Icarus*, **46**, pp.293-327, 1981.
15. R. Bjoerk, "Exploring the Galaxy using space probes", *International Journal of Astrobiology*, **6**, pp.89-93, 2007.
16. B. Golden and A. Assad (Eds.), "Vehicle Routing: Methods and Studies", in *Studies in Management Science and Systems*, **Vol. 16**, North-Holland, Amsterdam, 1988.
17. P. Toth and D. Vigo, "*The Vehicle Routing Problem*", Society for Industtrial and Applied Mathematics, Philadelphia, USA, 2001.
18. C. Cotta and A. Morales, "A Computational Analysis of Galactic Exploration with Space Probes: Implications for the Fermi Paradox", *JBIS*, **62**, pp.82-88, 2009.
19. G.A. Gurzadyan, "*Theory of interplanetary flights*", Gordon and Breach, Amsterdam, 1996.
20. E.J. Smith, D.E. Page, and K.-P. Wenzel, "ULYSSES - A journey above the sun's poles", *EoS*, **72**, pp.241-248, 1991.
21. V.G. Surdin, "Launch of a galactic probe using a multiple perturbation maneuver", *Astronomicheskii Vestnik*, **19**, pp.354-358, 1986.
22. http://voyager.jpl.nasa.gov/mission/weekly-reports/index.htm. (Last Accessed 3 June 2013)
23. F.J. Dyson, "Gravitational Machines", in *Interstellar Communication*, A.G.W. Cameron (ed.), Benjamin Press, 1963.
24. I.A. Crawford, "Some Thoughts on the Implications of Faster-Than-Light Interstellar Space Travel", *Quarterly Journal of the Royal Astronomical Society*, **36**, pp.205-218, 1995.
25. www.oica.net. (Last Accessed 3 June 2013)

(Received 10 May 2012; Revision & Accepted 10 December 2012)

* * *

JBIS, Vol. 66, pp.178-189, 2013

INNOVATIVE APPROACHES TO FUEL-AIR MIXING AND COMBUSTION IN AIRBREATHING HYPERSONIC ENGINES

CHRISTOPHER MACLEOD
School of Engineering, The Robert Gordon University, Aberdeen, Scotland, AB10 1FR, UK.

This paper describes some innovative methods for achieving enhanced fuel-air mixing and combustion in Scramjet-like spaceplane engines. A multimodal approach to the problem is discussed; this involves using several concurrent methods of forced mixing. The paper concentrates on Electromagnetic Activation (EMA) and Electrostatic Attraction as suitable techniques for this purpose - although several other potential methods are also discussed. Previously published empirical data is used to draw conclusions about the likely effectiveness of the system and possible engine topologies are outlined.

Keywords: Scramjet, mixing, EMA, controlled compression mixing, flow driven fuel system

1. INTRODUCTION

It is now around sixty years since the first serious attempts to produce a Scramjet engine. The progress made during the intervening time has been intermittent. Although new techniques like Computational Fluid Dynamics (CFD) mean that more is known about some of the detailed questions of aerodynamics, the current designs are little different to those suggested two generations ago. In recent years, programmes like HyShot, HyCause and the X-43A have again added to our knowledge but, despite optimistic claims from vested interests, extended flight under Scramjet power remains almost just as elusive as in the past.

There are several reasons why this is the case, some are due to the difficulty of designing the inlet and exhaust topologies for the extended flight envelope and similar mechanical-engineering concerns. However, although such issues are demanding, they have proved soluble in other aerospace applications and the most challenging aspect of Scramjet technology lies in fuel mixing and combustion.

In the high-drag and high-temperature regime where Scramjets operate, it is difficult to add further kinetic energy to an already excited flow-stream. This means that the engine is operating in a finely balanced region in terms of its thrust and drag and good conversion of the fuel's chemical energy into usable flow-energy is essential. However, at hypersonic speeds, air passes through the engine in around a millisecond, meaning that the fuel must mix with the air, burn and release its energy in a few tens of microseconds [1]. To achieve maximum extraction of energy, the fuel must be mixed stoichiometrically with air, at the molecular level, during this time. These operations should be performed in a way which does not disrupt the flow enough to cause an increase in drag. The resulting mixture has also to be burnt without the aid of the flameholding structures used at lower speeds - as projections into the duct would cause form-drag. Such considerations make it obvious why the technology is on the edge of practicality [2].

This paper builds on previous work [3, 4], published in *JBIS*, to suggest some engine topologies which might be used as a basis for experimentation or simulation into overcoming the mixing problem. For reasons discussed in the paragraphs below, these are not final design solutions, but are meant as a basis for discussion on Scramjet topologies and related issues.

In the work presented here, it is suggested that the answer to the mixing problem might lie in a multimodal system – that is, a system in which several different and complementary techniques are employed in order to achieve the mixing goal. In particular, the paper outlines how previous work on Electro-Magnetic Activation (EMA) and Electrostatic Attraction could be used to enhance and control the mixing and ignition process. Other potential methods that might be used in a multimodal system, including a new "flow-driven" concept, are also discussed.

The fuel-air mixing system in a scramjet engine is difficult to simulate accurately using techniques like CFD. Some of the physical phenomena present are hard to study - because of the speeds and temperatures involved, and are therefore rather poorly understood. Others, because they represent elements of both non-continuum (free molecular) and continuum (Navier-Stokes) flow or complex systems of interacting flows, are not readily amenable to standard modelling equations or simulation techniques. This is particularly true of the methods outlined in this paper, because they themselves are innovative in their approach and use unusual topologies and techniques to address the mixing problem.

Although modelling and simulation are difficult for the reasons outlined above, quite a number of authors have published their observations on applicable systems and these can be used to build useful conclusions. Therefore, the approach taken here is to base some of the discussion on previously published empirical and experimental results. It is for this reason, as already mentioned, that the suggestions presented are not final design solutions, but are meant to stimulate ideas for discussion. However, it is hoped that they can form the basis of future experiments or simulations to establish their credibility.

The paper deals with the dynamics of the fuel-air mixing

and combustion system and not with the aerodynamic design of the ducts or surfaces of the engine. This is because there are already many published analyses of these [2, 5]. Where it is necessary to state aerodynamic or flow parameters, figures from a well known design by Billig [6], discussed in Anderson [7], will be used. This design and its associated flow parameters are typical of the available examples and have been used in many previously published papers.

The discussion starts with a description of traditional mixing in scram systems and then goes onto discuss EMA assisted mixing, electrostatically assisted mixing, other multimodal approaches and finally some possible engine topologies.

2. THE CONVENTIONAL VIEW OF MIXING

In order to understand issues with the mixing process and how it may be improved, it is first necessary to understand its dynamics in more detail.

There are three simplified cases which illustrate the basic mechanisms of unforced mixing. The first is zero-shear mixing, which assumes invisid flow. In this circumstance, the fuel and air mix by simple diffusion only - there are no shear forces or macroscopic mixing due to turbulent flow.

Although this type of mixing is not practical due to its slowness, it is nevertheless an important case to understand. This is because, as mentioned in the previous section, there needs to be stoichiometric mixing at the molecular level for good energy extraction - and only diffusion (simple or forced) can supply this. So, no-matter how much the fuel is enfolded in the air by turbulence or other macroscopic mechanisms, the two components still need to diffuse into each other. Simple diffusion is controlled by Fick's Law, the form of which can be found in any textbook on the subject [8]. The special case of diffusion of two gases is covered in reference [9]. The topology usually considered is shown in Fig. 1.

The mixing layer, of fully diffused gas, thickness δ, grows down-stream [2] and has the approximate width:

$$\delta \approx 8\sqrt{\frac{Dx}{u}} \tag{1}$$

Where D is the molecular diffusivity for air and the fuel, x is the distance along the axis of the duct from the point of contact and u is the average of the speed of the air and fuel. The air and fuel will be completely mixed when δ is equal to the width of the duct b ($= b_1 + b_2$ in Fig. 1). The length along x required to fulfil this condition (denoted L) is:

$$L = \frac{ub^2}{16D} \tag{2}$$

Working out some practical figures for required mixing lengths and fuel velocities shows that L/b is of the order of 1500. Such long engines are not practical because of the associated skin-friction losses [5].

In the second mixing case, viscous laminar interaction is allowed - this is called the Laminar Shear Case [10], here:

$$\delta = 8\sqrt{\frac{\nu x}{u}} \tag{3}$$

Where ν is the kinetic viscosity (μ/ρ). Perhaps surprisingly, this produces no great improvement over the zero-shear case - because lateral movement is still by molecular processes.

In the final case, by increasing the velocity difference between the two streams or by other similar means, Kelvin-Helmholtz Instability may be induced and hence turbulent mixing. Here mixing is much more efficient because the generated vortexes entrain the fuel and air components and allow for contact at small scales. It is difficult to derive theoretical models for this, but several empirical surveys do exist. One useful and oft-cited treatment of turbulent mixing [11] showed that:

$$\delta = C\left(\frac{1-r}{1+r}\right)x \tag{4}$$

Where C is a constant reported to be between 0.25 and 0.45 in various experiments, and r is the ratio of the velocities of fuel and air.

Inserting some typical figures into this equation shows that turbulent mixing is much more efficient at producing mixing in a fairly short length of duct. However, the induction of instabilities in the flow is a macroscopic process - therefore slow and, in any case, diffusion is still necessary across the entrained fluid boundaries to achieve the molecular mixing necessary for efficient combustion. These factors mean that the achievement of a true stoichiometric mixture, in the time allowable, is still very difficult.

To induce turbulent mixing of the type described above, normal injection solutions (that is, fuel is injected "normally" or perpendicularly to the airflow) have been introduced as shown in Fig. 2.

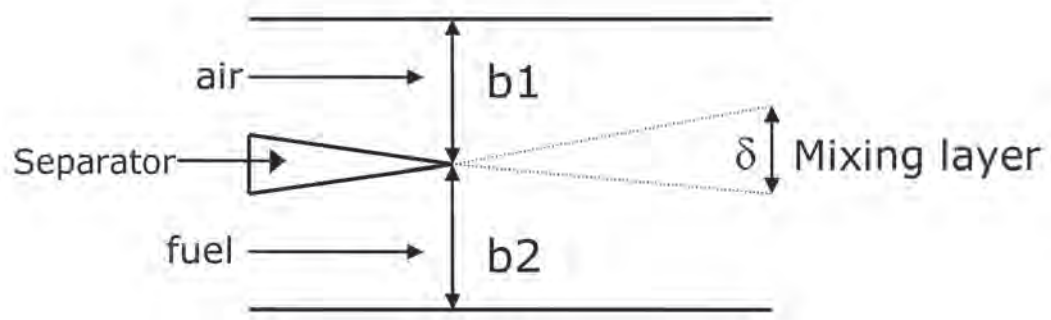


Fig. 1 Topology for zero-shear mixing.

Fig. 2 Normal and ramp injection.

One obvious problem with this is that such injection necessarily causes disruption to the flow and therefore strong shock interactions. These result in further drag - particularly if the boundary-layer is disrupted, as it inevitably will be. Further, this will inhibit mixing for the reasons discussed in the next section. Unfortunate inlet to combustor interactions may also upset the duct pressure balance and cause unstart [1]. Hence, even without considering compressibility issues, there are two conflicting requirements on the system, firstly good mixing but secondly minimum disruption to the flow.

The scale of the problem presented above can be seen by using some typical figures available for practical flow-speeds, fuels and engine lengths; for example, using Billig's engine design [6]. In this case, given that the maximum diffusion coefficient of hydrogen into air is around 2.6 $cm^2\ s^{-1}$ [8] (just below its auto-ignition temperature with air, at a typical Scramjet combustor pressure [6]), then the air and hydrogen must be turbulently mixed or entrained over the entire volume to a contact dimension of between a few millimetres and a centimetre if diffusion is going to act within the time available for complete mixing at high Mach [9]. Although this is a very difficult engineering task, it is not impossible to envisage that a viable Scramjet can be achieved. However, the results from actual scramjet tests do not show the predicted power output, and therefore something is wrong with the model. The most likely reason for this is discussed in the next section.

3. COMPRESSIBILITY AND MIXING

Results in a number of published papers demonstrate why the simple theoretical treatment of mixing presented in the last section needs revision. These results revolve around issues involving the relative speed (and therefore compressibility) of the two mixing flows (the air and fuel). They show quite clearly that, if the flows are relatively supersonic (as they would be in many practical mixing scenarios, like normal injection), the mixing layer is much smaller than if they are subsonic relative to each other.

Some of the important papers in this field are by Papamoschou and Roshko who, in their best known paper [12], present an extensive set of experiments and a theoretical framework. Azim and Islam [13] also carry out similar experimental work and again find that (to quote their abstract): "The mixing layer growth was found to decrease with increasing velocity ratio." Earlier work which also supports these findings includes Birch and Eggers [14] and Brown and Roshko [15], among others.

Although there is much debate about a theoretical framework to support such results, they certainly seem to be a product of compressibility phenomena. Consider, for example, a subsonic flow and one which is relatively supersonic; where these two flows meet, a shockwave forms. Although the shockwave is thin, it still represents a high density discontinuity between the flows. This region is maintained only by the energy flowing into the system, as this is required in order for the flow molecules to maintain their positions in equilibrium. The forces on the individual molecules have the form shown in Fig. 3 [16].

To the molecules on the lower energy side of the shock, the shock region is effectively a barrier to penetration (and therefore diffusion) - as to move into it would mean moving, against a steep energy gradient, into the area labelled "A" in Fig. 3. Even before shockwaves form, a region of increased compression exists which can have a similar effect. To quantify the amount of compression at the boundary between the flows, many authors define a relative speed for the flow components (essentially shifting the frame of reference from the laboratory to that of the free flow). This is often termed the convective Mach number *Mc*. A common definition for two flows is:

$$M_c = \frac{u_1 - u_2}{a_1 + a_2} \tag{5}$$

Where u_1 and u_2 are the speeds of the flows under consideration and a_1 and a_2 are the speed of sound in these flows.

In a theoretical and review paper [17] Slessor and his co-

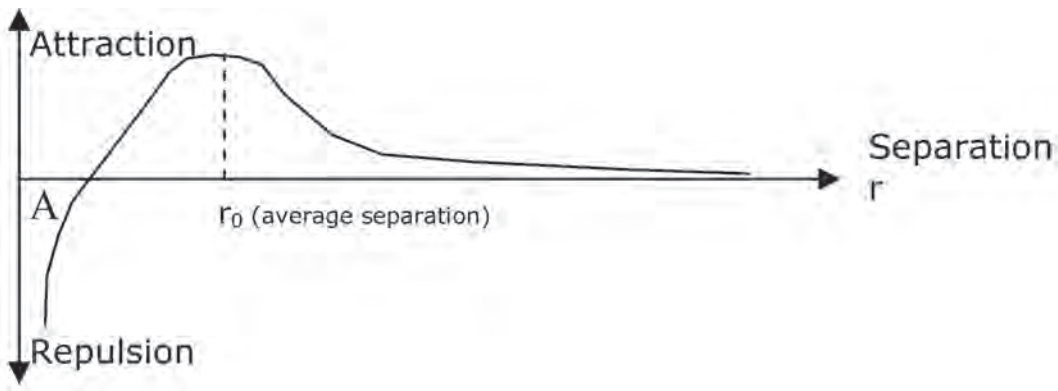


Fig. 3 Forces on flow molecules.

authors present a survey and consolidation of the results from ten previous experimental papers published between 1966 and 1998. Figure 4 shows a graph plotted from this data. In this graph, the spatial rate of increase of the shear layer thickness is labelled δ' and the growth rate of the layer at $M_c = 0$ is labelled δ'_0 (sometimes called the incompressible growth rate):

$$\delta' = \left.\frac{d\delta}{dx}\right|_{\delta = f(M_c)} \quad \text{and} \quad \delta'_0 = \left.\frac{d\delta}{dx}\right|_{M_c = 0} \tag{6}$$

So the term δ'/δ'_0 is the rate of change of the mixing layer, normalised to the incompressible case.

It can be seen from the graph that the amount of mixing decreases rapidly with increasing *Mc*. It also appears to tend asymptotically to a value of around $\delta'/\delta'_0 \approx 0.2$, particularly for $M_c > 1$.

There are several issues with this interpretation of the data, however. Firstly, most of these experiments measure the size of the shear or turbulent layer rather than the mixing layer per say. This means that the figures should not be taken as measure of the extent of an even approximately stoichiometric mixture – this would be much more difficult to establish and almost impossible to measure at realistic speeds. A second issue is that, although Slessor *et al* have correlated the data from a number of experiments, there are still only around forty data points available. The high cost and specialised equipment needed for such investigations mean that this will be the case for the foreseeable future. Finally, not all the available data is consistent. Several sets show distributions which indicate that maximum rate growth is not at *Mc* = 0. For example, the data measured by Hall et al [18], is typical of the shape of these curves, as shown in Fig. 5 (on the same scale as Fig. 4). Such data usually shows a peak in the region $0.2 < Mc < 0.8$ and this may be related to the triggering of turbulent mixing in such setups.

Slessor and his co-workers acknowledge such results and point-out they tend to occur at large density or velocity ratios (which is important, since one or both these conditions usually apply in the case of fuel-air mixing). They propose a new measure of compressibility (labelled Π_c), to integrate more of the outlying data points into an overall expression:

$$\Pi_c = \frac{\sqrt{\gamma_{(\max)} - 1}}{a_{\max}}\left(u_1 - u_2\right) \tag{7}$$

Where $\gamma_{(\max)}$ is the maximum specific heat ratio between the flows under consideration (there may be more than two) and similarly $a_{\max}$ is the maximum speed of sound of the components. However, as the authors themselves point out, not all the data fits even this formulation.

As discussed in the paragraphs above, there are problems with the available experimental data. These may be summarised as: lack of agreement on the theoretical basis of the effects observed, a general lack of experimental data and data sets which do not sit well with proposed "best fit" lines and curves. However, despite this, it is clear that, to a first approximation, mixing efficiency generally decreases with increasing compression (velocity ratio) between flow components. It is therefore essential that any experimental mixing system provides the ability to supply the fuel so that performance can by maximised. This is a technically difficult task. The next section discusses how it might be achieved using a technique previously published in *JBIS – Electromagnetic Activation* or *EMA*.

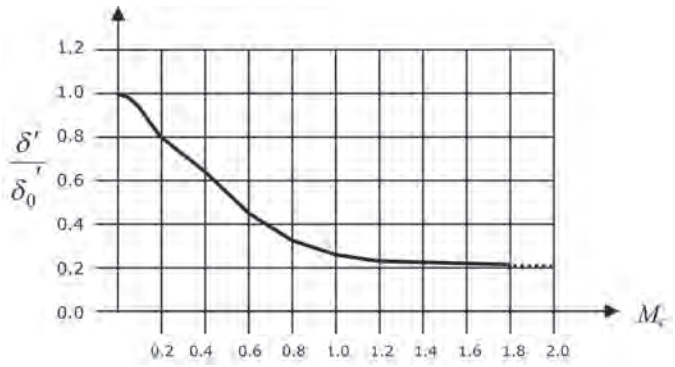


Fig. 4 Graph generated from data presented in Slessor *et al* [17].

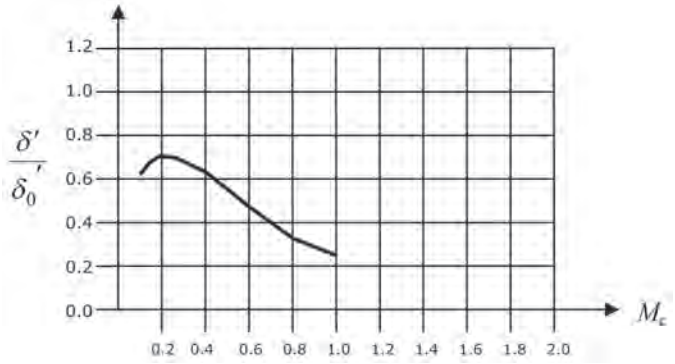


Fig. 5 Graph generated from data presented in Hall *et al* [18].

4. EMA ENHANCED MIXING

As discussed in the previous section, there is considerable debate over some of the physical mechanisms involved in mixing (and some of the more subtle aspects involved were also neglected in the description above for simplicity). However, the data clearly shows that the difference between optimal and worst-case laminar mixing efficiency is up to 500%. It is therefore imperative that the system supplying fuel to the airflow is able to provide it in a way which maximises the mixing effectiveness. In a real system, this means that (primarily) the speed must be controlled - but also (secondarily) the temperature, density and interaction angle. The EMA system discussed below is potentially capable of providing this flexibility (both in terms of experimental and final systems).

The principle idea behind EMA is that a fluid (either in a flow or at rest) is heated by Electromagnetic Radiation emitted into the fluid and absorbed by it. This radiation directly activates the fluid at the molecular level, increasing its internal energy by coupling to its rotational, translational, vibrational or electronic energy modes. A detailed theoretical discussion of this, together with derivations of the fundamental equations, is contained in previous papers [3, 4].

There are several ways in which EMA can be used in propulsion concepts. These include: Firstly, heating the airflow directly using millimetre-wave frequencies or ultraviolet light. This possibility was discussed extensively in the original papers [3, 4]. Secondly, adding a substance with high absorbance to the working fluid. This heats-up efficiently and transfers energy to the main flow (the components of pure air normally have a

rather low absorbance). Thirdly, heating the flow so much that some of its components disassociate and high-energy chemical reactions can occur, which add further energy (for example, in air, reactions involving the nitrogen components). Finally, the case being considered here: using EMA to accelerate or activate the fuel (or flow) in a more conventional scramjet type system with the aim of increasing its mixing efficiency.

The original papers mainly discussed adding heat to flow through open ducts; however, exactly the same principle can be applied to closed ducts like chambered de-Laval nozzles, as shown in Fig. 6.

As shown in the original EMA papers, the path-length x required to absorb a particular proportion of the radiation is given by:

$$x = \frac{\ln\left(1 - \frac{\varpi}{100}\right)}{-(\sigma n)} \tag{8}$$

Where ϖ is the percentage of radiation to be absorbed; for example, if this figure was 99%, then x would be the distance required to absorb 99% of the radiation power. The symbol σ is the Absorption Cross Section of the species involved, usually specified in cm^2 (in which case, x is in cm) and n is the number density of the species in particles per cm^3.

The radiation propagates in the duct by wall or mirror reflection in a similar way to propagation along a waveguide. The required mirror length D of a parallel sided structure, like that shown in the figure, is given by:

$$D = x \sin\theta \tag{9}$$

Where θ is the launch angle measured from a normal to the duct wall. The temperature rise of the fluid, in the radiation field, for a given mass flow-rate and absorbed power is:

$$\Delta T = \frac{\xi}{\dot{m} C} \tag{10}$$

Where C is the Specific Heat Capacity of the gas and ξ is the absorbed radiation power. The mass flow-rate may be calculated from $Av\rho$ (where A is the cross-sectional area of the chamber). Knowing these figures, it is then possible to calculate the exist velocity of a nozzle with choked flow:

$$v_e = \sqrt{\frac{2T_c R\gamma}{w(\gamma - 1)}\left[1 - \left(\frac{p_e}{p_c}\right)^{\frac{\gamma-1}{\gamma}}\right]} \tag{11}$$

Where T_c is the chamber temperature, w is the molecular weight of the working fluid, R is the gas constant and p_e and p_c are the pressures in the chamber and at the exit of the nozzle.

One important aspect of using radiation to heat the flow like this is its controllability and flexibility. The power applied to the fluid can be regulated electronically and although the example above illustrates a topology similar to a conventional rocket motor, this need not be the case. Because the heating effect may be applied at any point in the system, the shape of the nozzle or duct can be carefully controlled to provide a flow of any density, speed and temperature. EMA therefore has the capability of producing a much more flexible result than other alternatives like high pressure guns or simple electric heating.

In general, there are three options for EMA activated fuels. The first is simply to activate the fuel directly. The main way of doing this is to use a fuel molecule with a dipole moment - these can generally be activated in the microwave or millimetre parts of the spectrum [19, 20]; alternatively the vibrational modes of the molecule can be activated in the infrared [21]. Using micro or millimetre waves may be preferable, because of the efficiency and flexibility of the available sources (discussed in detail in [4]). Many hydrocarbons and other molecules can be activated like this [22] and The American Institute of Standards and Technology maintains on-line records of microwave absorption bands in the Hydrocarbon Spectral Database at its Physical Measurement Laboratory. Unfortunately however, many important simpler substances, like H_2 and CH_4, have no intrinsic rotational moment by virtue of their molecular symmetry.

The second option is to create or induce a dipole in an unpolarised fuel molecule. The ways of doing this are to either apply an external electric field (which acts to skew the shell electronic field and produce a polarised result); or alternatively, the electronic configuration of the molecule can be changed (usually by ionising it). Adding energy in this way has been explored in hydrazine thrusters [23].

The final option is to mix a high absorption substance in with the fuel - so that this heats up and transfers its energy to the non-absorbing fuel. As a simple example, consider water.

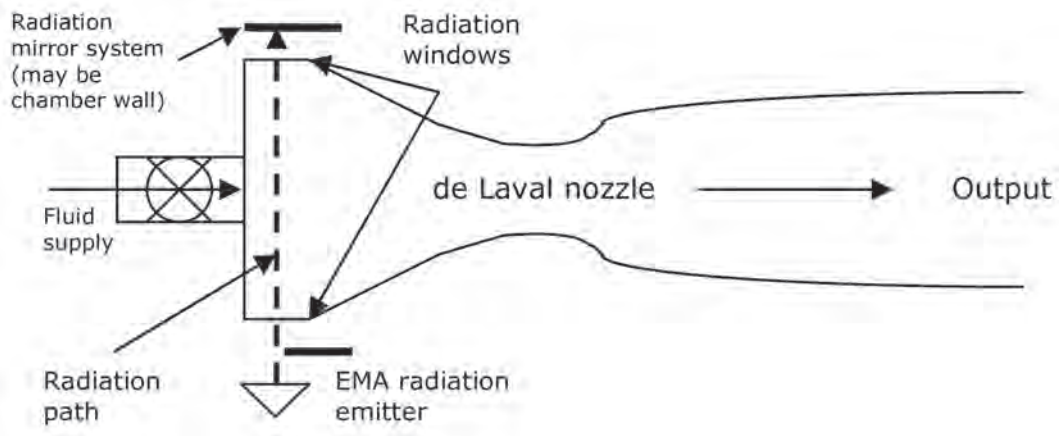


Fig. 6 Rocket-like EMA system.

This absorbs at several microwave and millimetre wave frequencies – for example 22 GHz, 183 GHz and 325 GHz [24]. In a atmosphere of H2 or a light hydrocarbon at Standard Temperature and Pressure (STP) where 1% of the molecules are water, half of the energy at 183 GHz will be absorbed in an 8 metre path-length and 92% in a 67 metre path-length; at 325 GHz almost all the energy is absorbed within a couple of centimetres.

Table 1 shows the power required to accelerate the fuel to the same speed as the airflow, using such an EMA system, in a Scram-like topology, based on Billig's figures for inlet conditions [6, 7]. The combustion output power is 100 MW (similar to a large turbojet) and the EMA system is assumed to be 40% efficient. The mixture ratio for H_2 is 1:4 by weight (as normally used in practical rocketry), rather then the stoichiometric 1:8; similarly for Kerosene, it is assumed to be 1:3.

From Table 1 it can be seen that the EMA system consumes between around 0.80% of the total power output (for H_2 at Mach 5) to nearly 60% (for kerosene at Mach 25). This illustrates the main issue with accelerating the fuel – the system is directing kinetic energy into the fuel electromagnetically, rather than thermally as in a rocket motor and so the overall efficiency will always be lower than in a topology where the fuel does not have to be speeded up. In a practical system, as much waste heat-energy as possible from the rocket cowl would be recaptured using heat exchangers to help power the EMA system.

The discussion above also assumes that the fuel feed is in-line with airflow. This is unlikely in reality as best mixing will probably occur at some other angle. Figure 7 shows the multiplying factor of the fuel velocity when subject to different angles of incidence with airflow (from 1 at 0° as just discussed to ∞ at 90°).

Using a system like this also raises several other possibilities. One of these is to chemically "engineer" fuels to have exactly the right properties. For example, a fuel with a useful dipole moment could be synthesised, as could one which flew apart during activation to produce lighter products, or a folded fuel with a reactive centre which only unfurled as it rotated (or conversely, two products which only became active when they react). Another interesting possibility is that, because the fuel and air stream are moving at similar speeds, the fuel could be added to the airflow with "swirl" (a rotational component) to aid mixing. This is usually ineffective in more conventional systems due to losses associated with the shock interactions between the fuel and airstream.

Overall, the system might have a topology similar to that shown in Fig. 8. This is not to scale, but is meant to illustrate a possible layout of the various components. The combustion and exhaust structure is illustrated here as an aerospike so that they operate in a rocket-like mode and avoid the need for

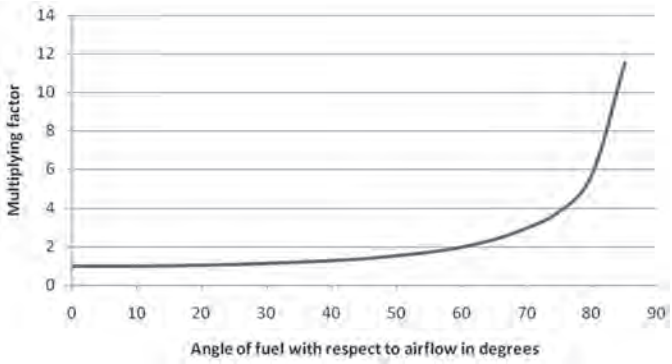


Fig. 7 Effect of angle of fuel injection with respect to airflow.

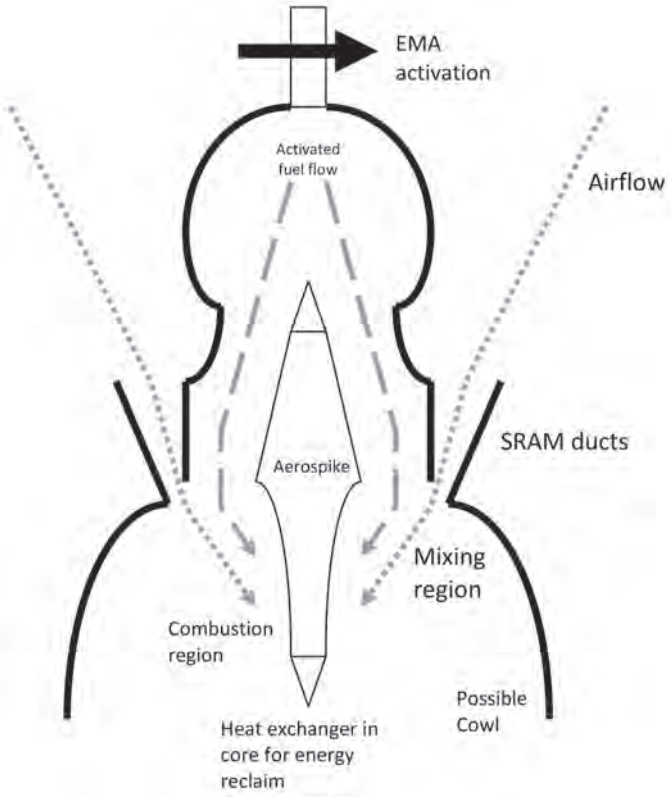


Fig. 8 Possible configuration topology.

flameholding structures (the EMA system is also capable of providing ignition energy if required). Such a structure also facilitates changeover between cycles in a combined-cycle system (for example, air breathing and conventional rocket).

5. ENHANCED MIXING USING FLOW-DRIVEN FUEL

The EMA system illustrated in the last section is not the only way to accelerate the fuel to match airflow speeds; electrical heating systems like arcjets, resistojets and others offer alternatives [25]. These are not reviewed here, as they are well

TABLE 1: *Power Consumed by EMA System at Various Speeds.*

Free stream Mach number	**5**	**10**	**15**	**20**	**25**
Speed at mixer inlet (ms^{-1})	561	1089	1650	2211	2739
EMA power into H_2 fuel system (MW)	0.785	2.65	6.77	12.0	18.75
EMA power into Kerosene (RP-1) System (MW)	2.351	8.24	20.28	36.0	56.26

covered in the literature - this discussion will instead focus on methods which have not been explored in detail elsewhere. One such idea draws the acceleration energy for the fuel more directly from the momentum (and therefore output power) of the engine - rather than an intermediate electrical system like EMA. Consider first a quantity of fuel at rest in a duct, as shown in Fig. 9a. If this is impinged on by a supersonic flow as shown in Fig. 9b, a shockwave passes through it and it almost instantly accelerates to essentially the same speed as the impinging flow, this is mediated by the viscosity of the gas and is normally accompanied by an increase in temperature, Fig. 9c.

This situation is extensively studied in shock tubes [26]. The air coming in from the left is known as the driver (or "pusher" in the discussion below) and the stationary fuel, the driven. Although the gases mix slightly at their interface [27], they remain essentially separate, providing that that their density ratio is low enough. Since the driven gas is initially stationary, there will be a strong interaction between it and the high-mach driver and so, to a first approximation, strong shock limits may be applied. In this case, the following relationships give the pressure, density and temperature ratios of the two gases:

$$\frac{p_2}{p_1} = \frac{2\gamma M}{\gamma+1} \quad \frac{\rho_2}{\rho_1} = \frac{\gamma+1}{\gamma-1} \quad \frac{T_2}{T_1} = \frac{(2\gamma^2+\gamma)M}{(\gamma+1)^2} \qquad (12)$$

Where the subscript 1 applies to the incoming gas and 2 applies to the stationary one. M is the mach number of the incoming flow (assuming, of course, as in the discussion above, a reference frame where the fuel is stationary).

To understand how this can be applied in a practical system, consider the following topology, shown in Fig. 10. The layout is similar to that for the EMA system shown in Fig. 8. The scram ducts, through which the main airflow passes, are permanently open. The fuel pusher ducts are opened and closed by a rotary valve, which may be attached to a rotating noise-cone. The cycle starts as shown Fig. 9a, with the injection of fuel when the valve is shut. When the valve opens it exposes the fuel to supersonic flow from the free-stream as shown in Fig. 9b. This accelerates the fuel as shown in Fig. 9c. It then exits into the main air flow at a similar speed. The cycle then repeats.

By having several pusher ducts which open, one at a time (or alternatively, symmetrically on oppose sides), a "Gatling-gun" type system could be set up, with some ducts closed and loading while others are open. In this way, fuel may be injected in a reasonably continuous cycle. A variation of this is to interlace the fuel-pusher and air ducts, so as to create a "sandwich" of air and fuel streams - this configuration is shown in Fig. 11.

This topology has some advantages over the EMA system outlined in the previous section. As stated earlier, the power to accelerate the fuel is derived mostly from the forward momentum of the engine – and therefore more directly from the combustion of the fuel (and so practical efficiency should be higher). Although power is still needed to turn the rotary valve and redirect the flow, this is equal to power-density difference between the two flow states and is much less than the power directly supplied to the EMA system. The nose-valve opening can be profiled to minimise the transient disruption to the flow as it moves.

There are also some disadvantages with this method. Firstly, it obviously introduces moving parts into the system and needs careful mechanical design. Secondly, the system is less controllable – there might be a particular problem with temperature of the fuel (unlike the EMA system, the fuel is in contact with an oxidiser as it is accelerating) – this makes the

Fig. 9 A stationary gas impinged on by a supersonic flow.

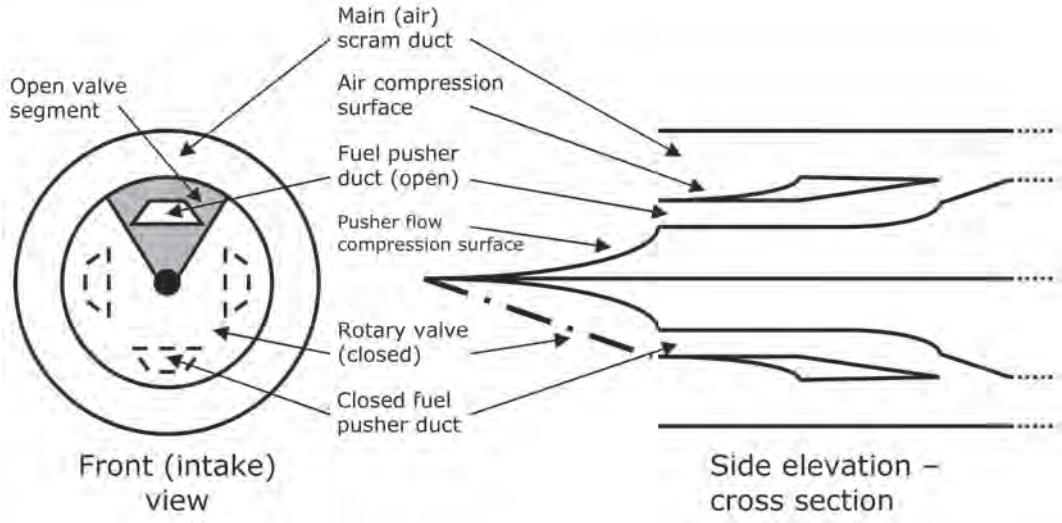


Fig. 10 Fuel pusher system using rotary valves.

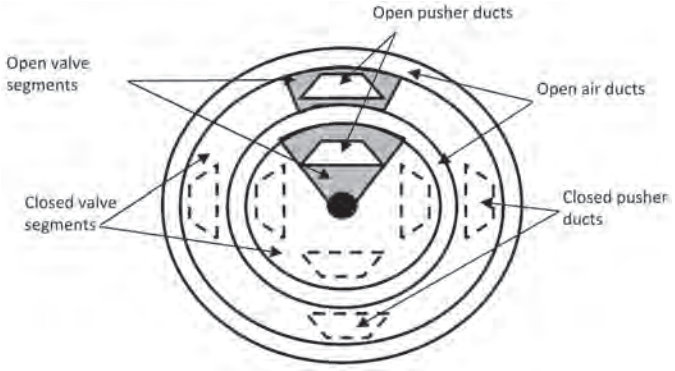


Fig. 11 Interlaced fuel and air topology.

design of the pusher-duct aerodynamics critical as their topology can control the temperature (temperature problems may also be overcome perhaps using pre-cooling of the fuel or duct, insulating pads or a buffer gas). Finally, the system is pulsed, which may cause vibration, stress or combustion problems.

There are many variations on the theme and possible combinations with the EMA method already described. For example, a system like this could be used to power the EMA device or provide mixing of the airstream with a radiation-absorbing gas. Alternatively, more than one gas could be used – for example: one pushing on the other and acting as an inert buffer, part of a two component fuel or a fuel premixed with a radiation-absorbing gas. Likewise, it may be possible to control the density ratio of the fuel and air so that the air mixes with, rather than drives, the fuel. However, this would probably require the fuel to be in a low density state (so that the flow was non-continuum and did not form a normal shock at the boundary).

6. ELECTROSTATIC ENHANCEMENT

One of the theses of this paper is that several different methods of forced or enhanced mixing may be used together to achieve good molecular penetration of the fuel into the air-stream. These approaches may be expected to be much more effective in the absence of the compressibility effects (and hence fluid-boundary shocks) discussed in the preceding sections. This approach was termed multimodal in the introduction. One potentially important method, which may be used with techniques already discussed, is electrostatically enhanced mixing. In this case, the fuel (or air) is charged and attracted into the air (or fuel) stream by a static charge as shown in Fig. 12.

In such a circumstances there are several forces acting on the fuel ions as shown in Fig. 13.

The electrostatic force is due to the field generated by the plate. Its magnitude is given by:

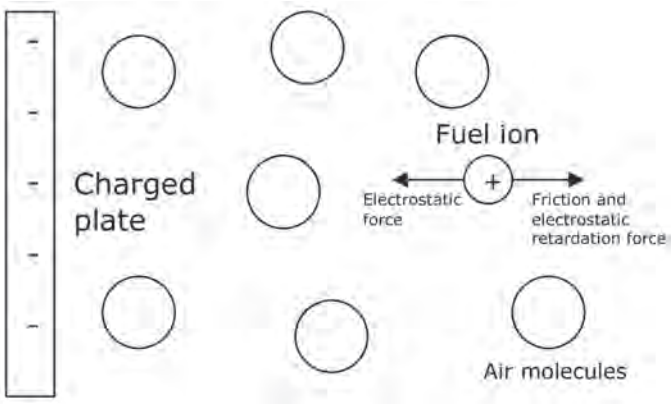


Fig. 13 Forces acting on a fuel ion.

$$F = qE \tag{13}$$

Where q is the ionic charge and E is the electric field magnitude. Expressions for E in various topologies are given in any relevant textbook [28]. The magnitude of the friction and retardation forces are complex, however the speed v of the ion can be simply expressed in terms of its mobility μ in air:

$$v = \mu E \tag{14}$$

The mobility varies with fluid parameters like density, temperature and composition and is also variable and non-linear at extremes of field intensity and viscosity. Many researchers have made measurements of ion mobility in air – from early workers [29, 30] to more recent attempts [31]. These values vary somewhat, but at the conditions of interest here, the minimum quoted values for small positive ions are around 1.15×10^{-4} $m^2V^{-1}s^{-1}$ (often quoted as 1.15 $cm^2V^{-1}s^{-1}$), a figure which is inversely proportional to air density [32]. Small negative ions typically have higher values. Using the available figures conservatively and assuming an electrical field of 0.5×10^6 Vm^{-1} (a sixth of the typical air breakdown field at STP $\approx 3 \times 10^6$ Vm^{-1}), the depth of the mixing layer, assuming electrostatic forcing only, may be evaluated as shown in Table 2.

These figures mean that electrostatic enhancement should approximately double (being conservative), the penetration of the fuel into the airflow, and perhaps increase it (being optimistic) by up to a factor of ten.

It might also be possible to increase the effect by using both air and fuel ions as shown in Fig. 14a. Other variations are also possible, for example modulating the charge on the plate spatially or temporally, Fig. 14b (a strong magnetic field would also achieve a similar effect). By moving the field generating potential to different positions along the plate, the ions could be directed into different positions of the airstream, so affording control which may be varied over

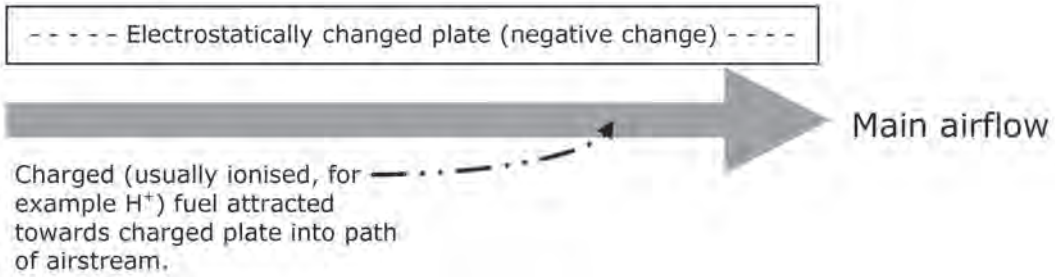


Fig. 12 The principle behind electrostatically enhanced mixing.

TABLE 2: *Penetration of Ionised Fuel into Air Stream Under Electrostatic Forcing Using Billig's Figures.*

Free stream Mach number	5	10	15	20	25
Speed at mixer inlet (ms^{-1})	561	1089	1650	2211	2739
Time taken for main flow to move 0.5m (μs)	891	459	303	226	183
Penetration of fuel in to stream in 0.5m stream movement (cm)	4.5	2.3	1.5	1.1	0.9

Figures are for small positive ions (very conservatively) $\mu = 1 \times 10^{-4}$ $m^2V^{-1}s^{-1}$ in a field of 0.5×10^6 Vm^{-1} medium is air at inlet to mixer in Billig's design [6, 7].

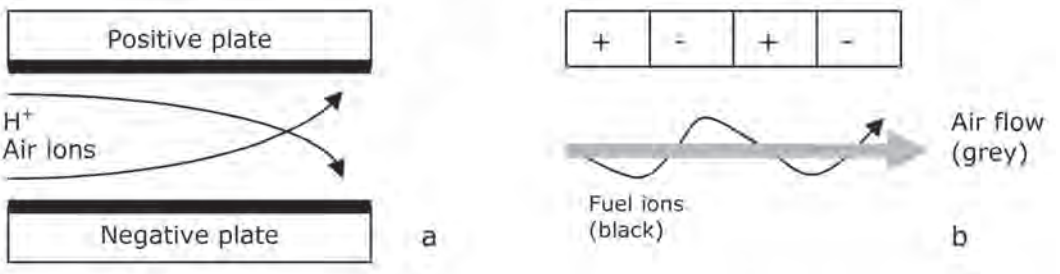


Fig. 14 Other field topologies.

the flight envelope. Fuel ions might also be accelerated along the engine axis as an alternative to the other methods already discussed.

In reality, since the fuel is initially confined, it should be much easier to ionise than the airflow [33]; indeed the EMA system itself could be designed to achieve this.

This system might be integrated into the previously discussed engine topology as shown in Fig. 15 (the central spike would be charged if the airflow were ionised).

7. OTHER APPROACHES

As well as the techniques discussed above, there are several other methods which could potentially help to enhance the mixing process. These are more speculative and so are only reviewed qualitatively.

The first such method is to use a diverter to force fuel into the airstream, Fig. 16. If the fuel is travelling at a similar speed to the airstream as previously discussed, the force from the diverter, which may include a shockwave boundary, should help to drive the fuel into the air.

Possible ways of implementing such a system are to use a rocket to generate the diverter stream – which, in turn, could be an EMA system, as already discussed or, alternatively, to use explosive diversion. This option is made more realistic by the recent development of new explosive materials which disintegrate producing fast molecular products [34, 35]. These could also be introduced into the fuel in pellet form.

Another potentially useful technique is to utilise Gas Dynamic Laser (GDL) effects. Because the composition of the fuel mixture can be controlled tightly in both the EMA and pusher topologies outlined above, inert gasses (particularly carbon dioxide) can be added to the fuel or generated in exhaust products. This allows a suitably shaped duct to generate a lasing effect [36] due to the thermally invoked population inversion. Such an effect could be used for a variety of purposes within the

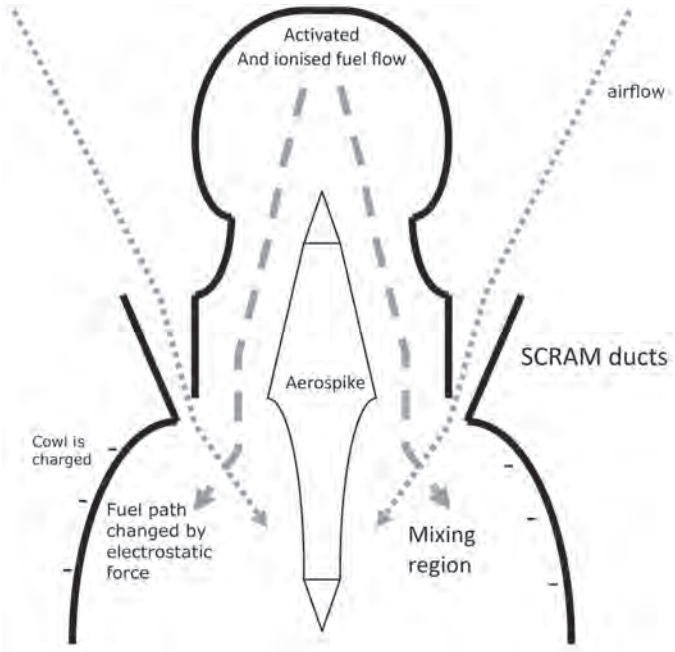


Fig. 15 Engine with added electrostatic forcing.

system. These include the extraction of energy from the stream in the form of (typically infrared) light. This energy may be used to power other subsystems, to remove heat from the flow (to produce a cooling effect) or to transfer it from one part of the flow to another (for example, from an inner duct to a bypass system). In turn, this might be used to aid mixing by increasing the internal energy in a controlled fashion or provide ignition.

8. PERFORMANCE PREDICTIONS AND ENGINE DESIGN VARIANTS

As explained in the sections above, there are several reasons why making predictions about mixing in hypersonic engines is difficult, and both the theoretical estimates and the measured results are subject to large uncertainties. This is particularly true of fuel injection and mixing techniques which produce complex flow patterns. One example of this is normal and steep-ramp

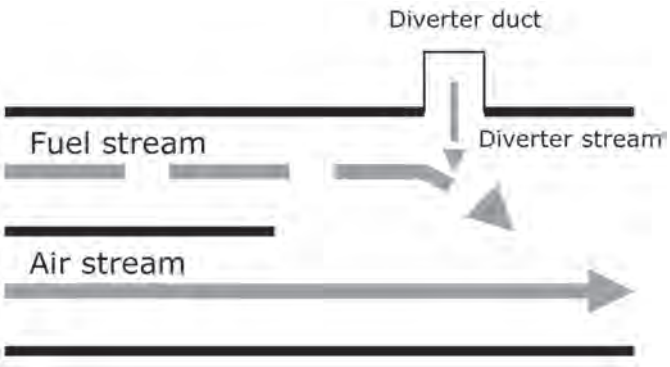


Fig. 16 Diverter operation.

injection (Fig. 2). In these cases, there is an initial normal or almost normal shock at the point where the fuel is released; this is a situation similar to the air-fuel interface at the inlet of the "pusher" topology discussed in section 5. Further downstream, the shock becomes oblique and eventually disappears as the fuel speeds up. Finally, the flow mixes turbulently in a similar way to that predicted by equation 4 (but, due to the shockwaves, well downstream of the initial contact region).

Although such issues make the mixing situation, using the techniques discussed in the previous sections, difficult to predict accurately, it is still possible to estimate performance. These predictions are based on the work presented in the earlier part of the paper and also on other published measurements and modelling of injection and mixing systems [37, 38], they are presented in Fig 17. Figure 17a shows predicted normalised mixing layer expansion enhancement. Here the symbols δ' and δ'_0 have the same meaning as in equation 6 and Fig. 4. Point A represents 1 unit and is the basic unforced diffusion-only situation of equations 1 and 2. Region B, which extends to a normalised mixing layer growth of 50 units, is the situation predicted in various papers, using equations similar to 3 and 4 and which does not factor-in the compressibility issues discussed in section 3. It is *very* dubious whether this is achievable in practice for the reasons already outlined (in other words, region B is probably nonexistent). Region C is the situation portrayed in regions A and B, but assuming that fuel is accelerated as discussed in sections 4 and 5; it extends from 5 units (for diffusive mixing only) to 55 units for good turbulent mixing. Region D, which extends from approximately 10 to 57 units, represents the addition of electrostatic enhancement, as discussed in section 6, applied to the basic diffusive mixing case of region C – the large range is due to uncertainly in parameters like ion-mobility, which are difficult to measure and not well characterised at high air and fuel temperatures and densities. Finally, region E stretches from approximately 100 to 550 units and assumes electrostatic enhancement, as in region D, and turbulent mixing. The graphs do not include the more speculative mechanisms outlined in section 7. The actual quantitative figures involved can be estimated from equations 1 to 4, the graph in figure 4 coupled with equations 5 and 6 and the information in Tables 1 and 2.

Figure 17b shows how mixing-layer growth decreases with increasing freestream Mach number; the figures are normalised to the situation at Mach 5 and based on Billig's aerodynamic figures. The decrease in mixing is mainly due to assuming a fixed mixing (longitudinal) length in the engine, but also on other changing parameters like density and increasing temperature. The gray region represents the uncertainty in the figures.

As well as the engine topology discussed earlier in the paper (for example as illustrated in Figs. 8, 10 and 15), other configurations are also possible. The benefits of some of these are discussed in the literature [2, 5]. Figure 18 illustrates two possibilities – diagram *a* shows an "inside out" configuration with the main airflow travelling centrally, and *b* an "interlaced" (sometimes called manifold [2]) configuration. The advantage of the first type is that it may prove less resistant to airflow; it might also be an easier way to produce a combined cycle (air-breathing at lower speeds, rocket-cycle at higher), than the other topologies. In the case of the second type, the purpose is to aid mixing by forcing more contact area between the fuel and the air.

There are many other design variants based on these themes – for example by allowing fuel injection from the centre (spike) of the engine or using Busemann topologies to reduce wave drag. In all these topologies, great care must be taken to engineer the aerodynamics of the duct so that shock-waves produced by the edges of the engine components do not interfere with the mixing process.

There may also be advantages to combining the interlaced configuration with the "pusher" topology discussed in section 5, and using the rotary valve to pulse *both* the air and fuel

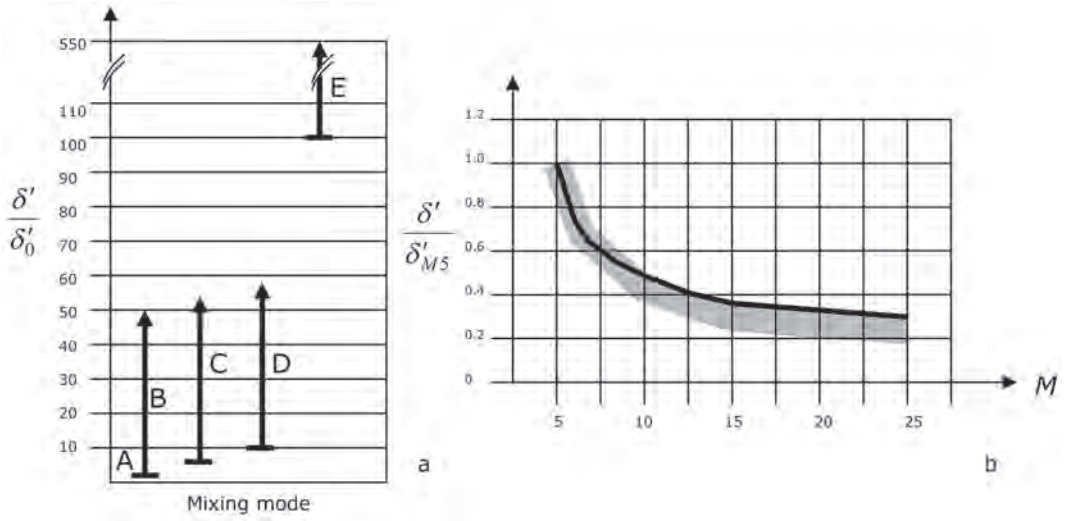


Fig. 17 Predictions of mixing performance based on the previously discussed models.

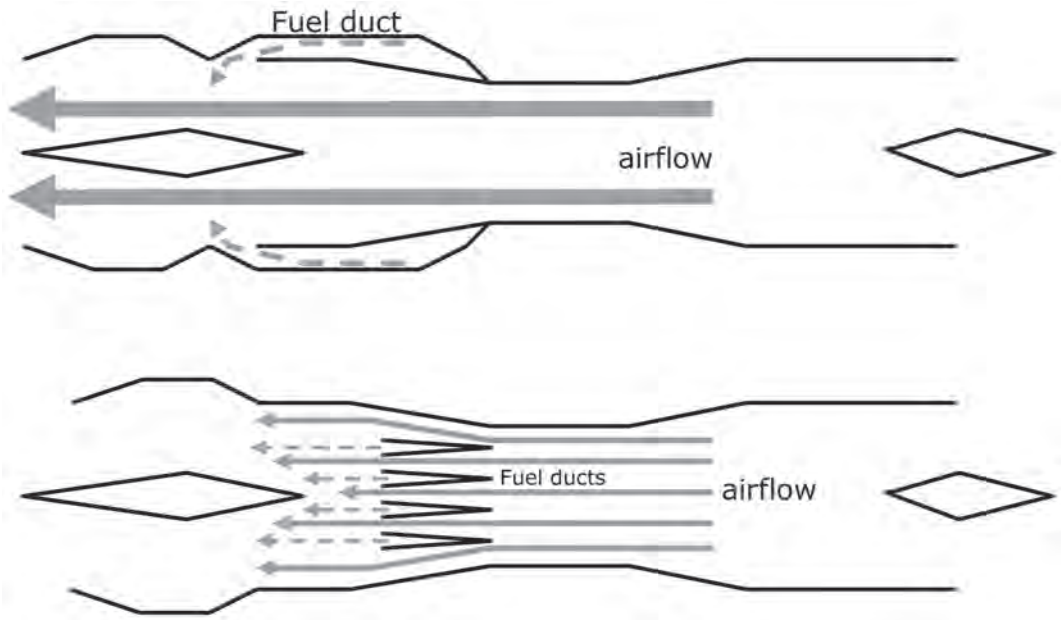


Fig. 18 Alternative engine topologies.

supply. This would produce a true sandwich of both species as shown in Fig. 19, without the edge effects caused by a pulsed fuel component penetrating a continuous air-stream.

If the fuel were premixed with a radiation absorbing gas, as mentioned in section 4, EMA could also be applied, in this case, *after* injection of the pushed fuel [3].

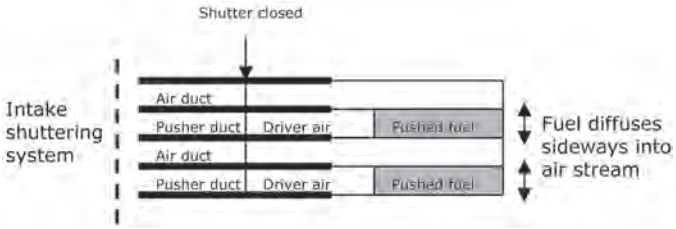


Fig. 19 Pulsing both air and (pushed) fuel.

9. CONCLUSIONS

The main aims of this paper are threefold. Firstly, to discuss the role of compressibility in mixing, provide a literature survey of some of the key papers in the area and illustrate why current engine designs do not live up to expectations. Secondly, to demonstrate that, although simple mixing regimes will probably not be able to provide the required performance, innovative forced mixing methods may be able to. Finally, to provide a variety of potential techniques that might lead to useful mixing enhancement and therefore a working system. The arguments presented in the sections above suggest that the potential improvement in mixing layer growth, using such techniques, is between 5 times (being conservative) and 550 times (being extremely optimistic), over simple injection. The most likely scenario is an improvement of between 10 and 100.

The methods outlined here are not the only ones which might be used to achieve enhanced mixing (for example, the injection and bursting of pressurised fuel capsules in the air-stream has not been discussed, but is another potentially interesting technique, as are innovative heat-exchangers, like those proposed in the SABRE engine [39]). Rather than provide a complete solution, they are meant to stimulate debate on the benefits of forced mixing and also of adapting a multimodal approach to the problem – that is, using several different but complementary methods to achieve a good mixture. Such an approach might be considered rather inelegant, but is probably necessary to overcome the considerable obstacles outlined in the first sections of the paper.

As discussed several times in the preceding sections, it should be remembered that many of the predictions described are subject of large uncertainties. Therefore, the next stage is to try and simulate some of the principles outlined using Computational Fluid Dynamics or similar methods and provide more accurate predictions of performance. However as outlined above, the different principles and forces at play and also the effects of compressibility make this a challenging task.

REFERENCES

1. J.P. Drummond, G.S. Diskin and A.D. Cutler, "Fuel-Air Mixing and Combustion in SCRAMJETS", *38th AIAA/ASME/SAE/ASEE Joint Propulsion Conference*, Indianapolis, Indiana, 7-10 July 2002. Paper AIAA-2002-3878.
2. W.H Heiser *et al*, "*Hypersonic Airbreathing Propulsion*", AIAA Education Series, Washington, 1994.
3. C. MacLeod, "Electromagnetically Activated Hypersonic Ducts", *JBIS*, **62**, pp.99-109, 2009.
4. C. MacLeod, K.S. Gow and N.F. Capanni, "Some Practical Aspects of Electromagnetic Activation", *JBIS*, **62**, pp.320-331, 2009.
5. E.T. Curran and S.N.B. Murthy, "Scramjet Propulsion", in *Progress in Astronautics and Aeronautics*, Series, **Vol. 189**, AIAA, Washington, 2001.
6. F.S. Billig, "Design and development of single stage to orbit vehicles", *John-Hopkins Applied Physics Laboratory Technical Digest*, **11**, pp.336-352, 1990.
7. J.D Anderson, "*Fundamentals of aerodynamics (3rd ed)*", McGraw-Hill, New York, 2001, pp.538-542.
8. R.B. Bird, W.E. Stewart and E.N. Lightfoot, "*Transport Phenomena*", Wiley, New York, pp 518-525, 2002.
9. Anon (Compressed Gas Association), "*Handbook of Compressed Gases (4th ed)*", Kluwer Academic Publishers, Norwell, pp 60-62, 1999.

10. H. Schlichting, "Boundary Layer Theory", McGraw-Hill, New York, 1979.
11. S.N.B. Murthy and E.T. Curran, "*High Speed Flight Propulsion Systems*", AIAA Progress in Astronautics and Aeronautics Series, **Vol. 137**, AIAA, Washington, 1991.
12. D. Papamoschou and A. Roshko, "The compressible turbulent shear layer: an experimental study," Journal of Fluid Mechanics, **197**, pp.453-477, 1988.
13. M.A. Azim and A. K. M. S. Islam, "Turbulent mixing layer from two nonparallel streams", *The Aeronautical Journal*, **107**, pp.241-248, 2003.
14. S.L. Birch and J.M. Eggers, "A critical review of the experimental data for developed free turbulent shear layers", NASA document SP-321, pp.943-949, 1973.
15. G.L. Brown and A. Roshko, "On density effects and large structures in turbulent mixing layers", *Journal of Fluid Mechanics*, **64**, pp.775-781, 1974.
16. J.F. Douglas, J.M. Gasiorek, J.A. Swaffield, "*Fluid Mechanics (4th ed)*", Pearson (Prentice-Hall), Harlow, 2001, pp.7-9.
17. M.D. Slessor, M. Zhuang and P.E. Dimotakis, "Turbulent shear-layer mixing: growth-rate compressibility scaling", *Journal of Fluid Mechanics*, **414**, pp.35-45, 2000.
18. J.L. Hall, P.E. Dimotakis and H. Rosemann, "Experiments in nonreacting compressible shear layers", *AIAA Journal*, **31,** pp.2247–2254, 1993.
19. T.M. Sugden and C.N. Kenney, "*Microwave Spectroscopy of Gases*", Van Nostrand, London, 1965.
20. W. Gordy, W.V. Smith and R.F. Trambarulo, "*Microwave Spectroscopy*", Wiley, New York, 1953.
21. B.H. Stuart, "*Infrared Spectroscopy: Fundamentals and applications*", Wiley-Blackwell, London, 2004.
22. S. Thorwirth *et al*, "Rotational Spectra of small PAHs", *The Astrophysical Journal*, **662**, pp.1309-1314, 2007.
23. S. Satori *et al*, "The development of a microwave engine", *27th International Electric Propulsion Conference*, Pasadena, California, 15-19 October 2001. Paper IEPC-01-224.
24. J.H. Van Vleck, "The absorption of microwaves by uncondensed water vapour", *Physical review*, **71**, pp.425-433, 1947.
25. M. Martinez-Sanchez and J.E. Pollard, "Spacecraft electric propulsion – an overview", *Journal of Propulsion and Power*, **14,** pp.688-699, 1998.
26. J.K. Wright, "*Shock tubes (Methuen's Monographs on Physical Subjects)*", Spottiswoode Ballantyne and Co, London, 1961.
27. R.V. Vasil'eva *et al*, "Turbulent mixing of driver and driven gases in a shock tube channel", *Journal of applied mechanics and technical physics*, **26**, pp.271-277, 1985.
28. D.J. Griffiths, "*Introduction to electrodynamics*", Pearson, New Jersey, 1998.
29. H.A. Erickson, "The mobility of argon and hydrogen ions in air", *Physical Review*, **26**, pp.465-468, 1925.
30. H.A. Erickson, "The mobility of acetylene ions in air", *Physical Review*, **28**, pp.372-373, 1926.
31. H. Tammet, "Reduction of air ion mobility to standard conditions", *Journal of Geophysical Research*, **103**, pp.13933-13937, 1998.
32. V.A. Mohnen, "Formation, nature and mobility of ions of atmospheric importance", in *Electrical Processes in Atmospheres*, eds. Dolezalek and Reiter, Steinkopff-Verlag, Darmstadt, 1977, pp.1-17.
33. C. MacLeod, N.F. Capanni, K.S. Gow, "Fuel Encapsulation for Inertial Electrostatic Confinement Nuclear Fusion Reactors", *JBIS*, **64**, pp.139-149, 2011.
34. J.D. Anderson, "*Gas Dynamic Lasers: An introduction*," Academic Press, New York, 1976.
35. J. Hogan, "Why bombers choose explosives that fall apart", *New Scientist*, **2484**, p.14, 2005.
36. F. Dubnikova *et al*, "Decomposition of Triacetone Triperoxide Is an Entropic Explosion", *Journal of the American Chemical Society*, **127**, pp.1146-1159, 2005.
37. D.W. Bogdanoff, "Advanced Injection and Mixing Techniques for Scramjet Combustors", *Journal of Propulsion and Power*, **10**, pp.183-190, 1994.
38. J.M. Seiner, S.M. Dash and D.C. Kenzakowski, "Historical Survey of Enhanced Mixing in SCRAMJET Engines", *Journal of Propulsion and Power*, **17**, pp.1273-1286, 2001.
39. J.J. Murray, C.M. Hempsell and A. Bond, "An Experimental Pre-cooler for Air-breathing Rocket Engines", *JBIS*, **54**, pp.199-209, 2001.

(Received 26 June 2013; Revision 5 June 2013; Accepted 21 June 2013)

* * *

JBIS, Vol. 66, pp.190-194, 2013

GRAVITATIONAL ASSIST VIA NEAR-SUN CHAOTIC TRAJECTORIES OF BINARY OBJECTS

JOSEPH L. BREEDEN
Prescient Models LLC, 300 Catron St., Suite B, Santa Fe, NM 87501, USA.
Email: breeden@prescientmodels.com

Using a process similar to the creation of hypervelocity stars, this paper proposes the use of binary objects (binary asteroid, spaceship + asteroid, etc.) for interstellar travel. Previous research has shown that binary star - single star interactions can cause high-velocity ejection of one member of the inbound binary. By selecting the correct chaotic trajectory, the same may be attainable for ejecting the chosen member of a binary object targeted as near to the sun as is survivable by electronics and/or crew. This paper will outline the basic process and compute the velocity that can be achieved under various orbital parameters via a conservation of energy calculation. It is shown via analogy to previously published calculations involving binary star - black hole interactions that trajectories may exist to achieve useful energy gain.

Keywords: Interstellar Travel, Chaos Theory, Binary Asteroids, Hypervelocity Stars

1. INTRODUCTION

Gravitational assist is well known in spaceflight [1]. Simple gravitational assist from close approaches to moving bodies, typically planets, is used extensively for current unmanned space probes. These trajectories usually include an Oberth manoeuvre [2] to increase the acceleration during gravitational assist by firing engines at closest approach. This paper discusses the possibility of another kind of gravitational assist achieved when two objects in mutual orbit make a close approach to a massive body.

The interaction between binary stars and single objects has been shown to allow for the ejection of one of the incoming objects at high velocity [3, 4]. The high stellar density of globular star clusters causes binary star - single star interactions to occur at a comparatively high rate. This interaction has been shown to cause the diameter of the core of globular star clusters to oscillate [5], sometimes chaotically [6], because of the ejection of stars as a heating source for the cluster.

Similarly, hypervelocity stars have been detected in the Milky Way galaxy [7]. Sixteen such stars are known to date. These are stars traveling at galactic escape velocities. The best explanation for these extreme velocities appears to be the interaction of a stellar binary with the Milky Way's central black hole [8].

This process is also suggested for how Triton was captured by Neptune [9]. An inbound binary object was tidally disrupted, capturing one element of the binary and ejecting the other. The same process has been proposed as a general mechanism for how planets can capture moons [10].

The question addressed here is whether binary objects can be used in a similar way for astronautics. By appropriately targeting a binary object toward a massive body, such as the Sun, the goal is to eject one of those objects with a significant gain in velocity while leaving the other object in a close captured orbit. This paper computes the energy that can be transferred to one of the objects in the binary via a close approach to the Sun, and the resulting velocity obtained. This process leverages the intrinsically chaotic nature of a three-body system, so desirable trajectories should be rare, if present. Leveraging the study by Hills [8] it is shown that trajectories resulting in ejection may, in fact, exist. However, this paper cannot address yet whether such trajectories are survivable.

2. THE BASIC TRAJECTORY

The key to the binary object gravitational assist manoeuvre is having a disposable object of significant mass. For purposes of discussion, assume that an asteroid in near-Earth-orbit (NEA) is employed. As of September 2011, 8,121 NEAs are known [11] with 828 of those having diameters of 1 km or greater. A recent discovery has shown that such objects exist in "Trojan" orbits [12] near the Earth, inhabiting the Earth-Sun Lagrangian points L4 and L5, just as happens with the gas giant planets. The first such object was found at roughly +60° relative to the Earth sharing roughly the same orbit [13].

Using such an NEO as in Fig. 1, imagine that some part of the object is hollowed out and used as the interstellar space ship as proposed by D. Cole in the 1960s while the remainder is to be discarded. The NEO's orbit would have to be modified to achieve a close encounter with the sun.

En route to the Sun, the object would be split into a binary, presumably by destroying a "waist" that was created earlier in the object. Upon close approach to the Sun, assuming just the right trajectory, the expendable part would be captured into solar orbit and the "spaceship" would be ejected at high velocity.

The following results would also follow from assuming that a man-made spaceship enters orbit about an appropriately sized object, such as an asteroid or sun-grazing comet, and travels as a binary to near-Sun interaction.

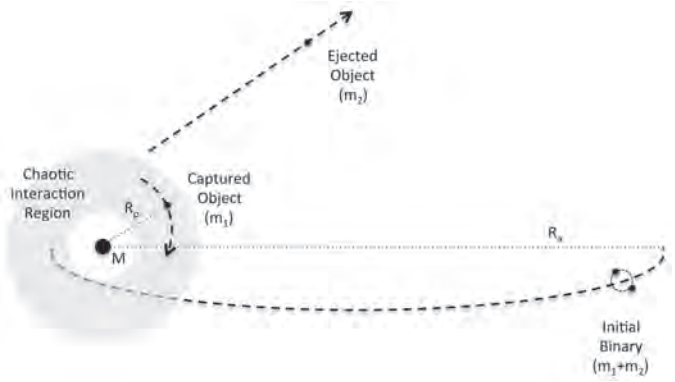


Fig. 1 Diagram of the binary-Sun interaction.

3. ENERGY TRANSFER CALCULATIONS

This calculation looks at how much energy can be transferred from the binary to one of the final parts via a close approach to the Sun. The primary source of energy gain is the change in gravitational potential energy between near-Earth orbit and perihelion. This is computed via a simple conservation of energy equation. Energy at aphelion equals energy at perihelion.

$$E_a = E_p \tag{1}$$

The initial energy will include orbital kinetic energy $E_{a,orbital}$, rotational energy of the binary $E_{rotational}$, and gravitational potential energy at aphelion U_a,

$$E_a = E_{a,orbital} + E_{rotational} + U_a \tag{2}$$

$$E_a = \frac{Gm_1m_2}{4R_b} + \frac{GMm_0}{2R_e} - \frac{GMm_0}{R_a} \tag{3}$$

where

- m_1 is the mass of the object that will be captured into solar orbit.
- m_2 is the mass of the object that will be ejected.
- $m_0 = m_1 + m_2$ is the total mass of the initial binary.
- M is the mass of the Sun.
- R_b is the separation of the binary objects' mutual orbit.
- R_e is the mean radius of the Earth's orbit about the Sun.
- R_a is the distance of the object from the Sun at aphelion.

The equation for E_a would not normally include a kinetic energy term at aphelion if this is purely an object in an elliptical orbit.

When the binary object reaches closest approach to the Sun, if the correct trajectory is attained, a transfer occurs that puts object 1 into close solar orbit with kinetic energy $E_{1\ p,orbital}$ and potential energy U_{1p} and object 2 is ejected with total energy E_2.

$$E_p = E_{1\ p,orbital} + U_{1p} + E_2 \tag{4}$$

$$E_p = \frac{GMm_1}{2R_p} - \frac{GMm_1}{R_p} + E_2 \tag{5}$$

where

- R_p is the mean orbital radius of object 1 attained via the transfer at perihelion.
- E_2 is the total energy of object 2 after the transfer at perihelion.

Equating energy at aphelion with energy at perihelion after the transfer, solve for $E2$ as

$$E_2 = G\left[\frac{m_1m_2}{4R_b} + M\left(-\frac{m_0}{R_a} + \frac{m_0}{2R_e} + \frac{m_1}{2R_p}\right)\right] \tag{6}$$

At infinity, the potential energy of object 2 will be zero, so the velocity of object 2 at infinity (v2,∞) is

$$E_2 = \frac{1}{2}m2v_{2,\infty}^2 \tag{7}$$

giving

$$v_{2,\infty} = \left\{2G\left[\frac{m_1}{4R_b} + M\left(-\frac{1}{R_a}\frac{m_0}{m_2} + \frac{1}{2R_e}\frac{m_0}{m_2} + \frac{1}{2R_p}\frac{m_1}{m_2}\right)\right]\right\}^{1/2} \tag{8}$$

4. RESULTS

Using Eq. 8 for the exit velocity of object 2, several things are immediately clear. The first term

$$\frac{m_1}{4R_b}$$

comes from the binding energy of the initial binary. That contribution is negligible compared to the other terms, which are all multiplied by the Sun's mass. In contexts such as binary star - single star interactions, the binary star binding energy can be a significant contributor or even the primary contributor. The same is not true here.

Assuming that perihelion will be much closer than the Earth's orbit, only the final term will contribute significantly to the exit velocity. Therefore,

$$v_{2,\infty} \approx \left\{\frac{GM}{R_p}\frac{m_1}{m_2}\right\}^{1/2} \tag{9}$$

Eq. 9 indicates that to maximize energy transfer, we want the smallest possible perihelion distance and for the "spaceship" mass m_2 to be much less than the mass of the deposited object m_1. Also note that the result depends very little on whether the original object's aphelion is greater than Earth's orbit even out to 100 AU, assuming perihelion is on the order of 1/4 AU.

In the calculations above, the binary object starts at aphelion - no radial velocity. For object 2, the velocity at infinity is given by $v_{2,\infty}$. Therefore, relative to the Sun, the net change in radial velocity is $\delta v = v_{2,\infty}$.

Using the full equation, we get the results in Fig. 2. We can approximate the travel time to Alpha Centauri assuming no additional propulsion as shown in Fig. 3.

When considering the possibility of sending unmanned probes to another star, more extreme trajectories are possible.

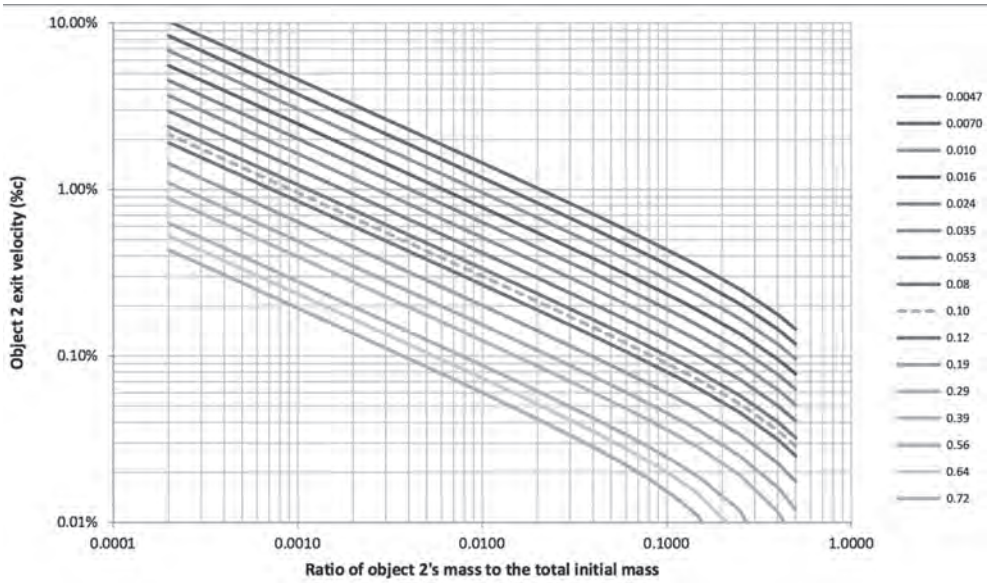


Fig. 2 The exit velocity of object 2 at infinity as a percentage of the speed of light, c, for a range of masses, m_2. The lines correspond to perihelion distance in AU including $\frac{1}{4}$ Mercury (0.10), Mercury (0.39), and Venus (0.72).

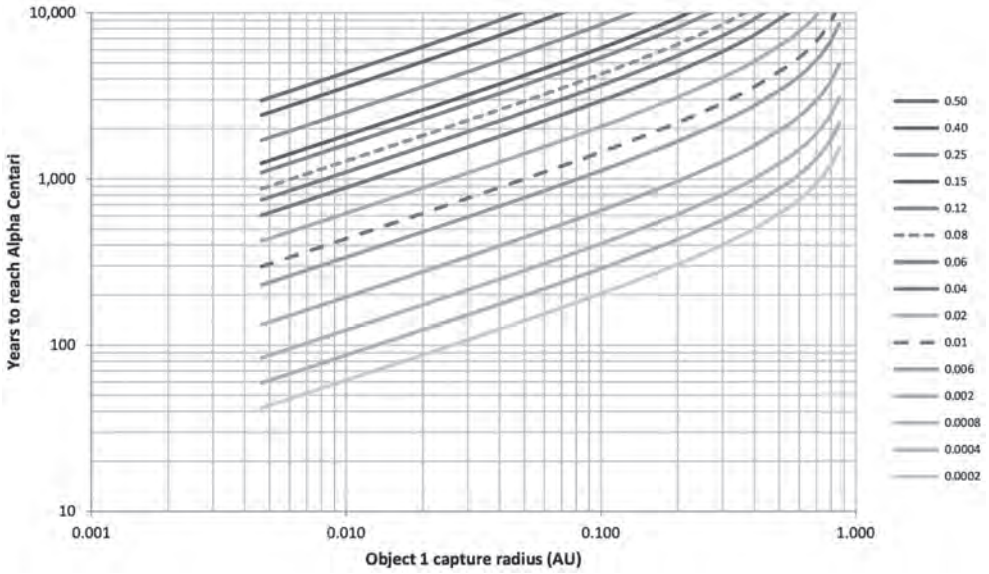


Fig. 3 The travel time needed to reach Alpha Centari as a function of the perihelion distance. The lines correspond to the ratio of m_2/m_0.

Current estimates are that electronics would survive a closest approach of 0.2 AU. If a usable trajectory could be found with perihelion of 0.2 AU for a probe with 0.1 mass ratio, the exit velocity would be 1,800 km/s or 0.06%c. With no other propulsion mechanism, that would correspond to a travel time of 7,300 years to Alpha Centauri. This result is highly sensitive to the perihelion distance and the mass ratio. Reducing the perihelion distance to 0.1 AU would increase the exit velocity by 50%. Conversely, increasing the mass ratio to 0.2 would cause an equivalent decrease in the exit velocity. Reducing the perihelion distance is largely an engineering question. The minimum mass ratio will be a physical constraint.

5. PROBABILITY OF EJECTION

The above calculation shows how much energy is potentially available under certain assumptions. It does not prove that trajectories exist that can extract this energy. Hills [8] conducted high-fidelity simulations of binary star - black hole interactions and measured the probability of an ejection. Hills computes a dimensionless parameter, D_{min}, as a function of the masses of the objects involved for which he shows the probability of an exchange. Replacing his black hole mass with a solar mass and his binary star mass with binary asteroid mass yields

$$D_{min} = \frac{R_p}{R_h}\left(\frac{2M}{10^6 m_0}\right)^{-1/2} \tag{10}$$

According to Hills' simulations, for $D_{min} \approx 1$, the probability of interaction across all randomly selected initial conditions is 1%. This probability drops roughly linearly to 0 for $D_{min} \approx 170$. This implies that in order to have a non-zero probability of asteroid ejection requires

$$\frac{R_p}{R_b} < 0.5 \tag{11}$$

In the current example, assuming a total mass for the binary of 10^{15} kg, which is approximately the mass of Deimos [14], a perihelion distance 0.2 AU, and a semi-major axis for the binary of 150 km, the probability of an interaction is approximately 0.4%. This means that a detailed search will be required, but that ejection trajectories are possible.

Hills' formula does not consider the mass ratio of the objects in the initial binary. However, for extreme mass ratios, ejection seems unlikely. This is unfortunate since extreme ratios lead to the greatest ejection velocity. Specific simulations will be needed to determine the range of mass ratios over which ejection can occur.

Hills computed several quantities in his follow-up work of 1991 [15] such as the average energy exchange and average ejection velocity. The objective here is to identify uniquely applicable extreme trajectories rather than the average results across all ejections, so Hills' average quantities are not directly applicable. Because of this and his use of equal mass objects, Hills' work and the adaptation above are suggestive of the possibility of ejection, but a specific search for extreme trajectories relevant to the current problem will be required to settle the question of whether usable trajectories exist.

6. SPECIFIC TRAJECTORIES

In addition to questions of velocity and probability of ejection, the details of specific escape trajectories will be important. In Fig. 4 of the 1983 paper by Hut and Bahcall [3], they show several examples of escape trajectories produced by three-body interactions. Certainly those examples were chosen to illustrate the inherently chaotic nature of the interactions, but no reason exists to believe that the extreme trajectories sought here will be any simpler. With chaotic trajectories, the G-forces that would be encountered by the ship and travellers must be considered carefully. It may be possible that the acceleration encountered as these velocities are achieved is too great for the survival of ship or crew.

Also, the margin for error in the computed trajectory is unknown. Even if the necessary extreme trajectories exist, are they navigable? At this point it must be assumed that the mutual orbit of the binary must be precisely timed to its close approach of the Sun in order to get maximum energy transfer and be propelled in the correct direction. One must assume that some form of propulsion will be needed for course correction, and that precision adjustments will be needed.

All of these questions could be resolved via simulations of specific orbits.

7. CONCLUSIONS

The process described above was designed to mimic what researchers believe may be occurring naturally across a mass range from binary star - black hole interactions to Kuiper belt object - gas giant planet interactions.

It also bears a resemblance in the energy calculations to the process of using a tether to accelerate an object [16, 17, 18] and to the two-burn Oberth manoeuvre [19, 20]. Although the energy calculations may be similar, the actual trajectories will likely be quite different. The chaotic trajectories creating acceleration here may be far more complex than the simple result of cutting a tether, although this is not necessarily a benefit.

For the technique described here, the most difficult question to answer is, "How is the energy transferred?" an analogy of dropping ballast from a balloon is tempting but inaccurate. Three-body interactions are well studied as chaotic systems where extreme results are possible, but it can be constructive to view this as a tidal disruption of the initial binary.

The calculations performed here show that trajectories may exist whereby a binary object-Sun gravitational assist manoeuvre could provide significant energy under plausible conditions. This could occur without the use of any significant propulsion in either engines or fuel. However, the most likely approach would incorporate elements of all known effects:

- Gravitational assist via a binary object-Sun interaction.
- Gravitational assist leveraging the proper motion of the Sun toward a nearby star like Alpha Centauri.
- An Oberth manoeuvre at perihelion following ejection.

Just as important as getting to another star is being able to stop once arrived. One option would be to perform this binary object-star gravitational assist again, but in reverse, as was hypothesized for the capture of moons by planets [9]. If the "spaceship" was a section of the original asteroid, one could envision splitting the remaining part again into a binary object and performing a close approach to the other star. With a true starship, it would be roughly equivalent to dropping ballast at exactly the right moment. In either case, upon reaching their destination, the crew will want to be in the portion captured into stellar orbit.

REFERENCES

1. M. Minovitch, "A method for determining interplanetary free-fall reconnaissance trajectories", Technical report, Jet Propulsion Laboratory Technical Memo TM-312 130, 23 August 1961.
2. H. Oberth, "Ways to spaceflight", Technical report, NASA TT F-622, 1929.
3. P. Hut and J.N. Bachall, "Binary-single star scattering. i - numerical experiments for equal masses", *Astrophysical Journal*, **268**, pp.319-341, 1983.
4. P. Hut, "Binary-single star scattering. ii - analytic approximations for high velocity", *Astrophysical Journal*, **268**, pp.342-355, 1983.
5. J.L. Breeden, H.N. Cohn, and P. Hut, "The onset of gravothermal oscillations in globular cluster evolution", *Astrophysical Journal*, **421**, pp.195-205, 1994.
6. J.L. Breeden and H.N. Cohn, "Chaos in core oscillations of globular clusters", *Astrophysical Journal*, **448**, pp.672-682, 1995.
7. W.R. Brown, M.J. Geller, S.J. Kenyon, and M.J. Kurtz, "Discovery of an unbound hypervelocity star in the milky way halo", *Astrophysical Journal*, **622**, pp.L33-L36, 2005.
8. J.G. Hills, "Hyper-velocity and tidal stars from binaries disrupted by a massive galactic black hole", *Nature*, **331**, pp.687-689, 1988.
9. C.B. Agnor and D.P. Hamilton, "Neptune's capture of its moon Triton in a binary-planet gravitational encounter", *Nature*, **441**, pp.192-194, 2006.
10. D.M. Williams, "Forming big moons through capture around gas giant planets", *Bulletin of the American Astronomical Society*, **42**, pp.1070, 2010.
11. NASA, "NEO discovery statistics", http://neo.jpl.nasa.gov/stats. (Last Accessed 3 July 2013)
12. P. Robutel and J. Souchay, "An introduction to the dynamics of trojan asteroids", in *Dynamics of Small Solar System Bodies and Exoplanets*, eds. R. Dvorak and J. Souchay, Springer, p.197, 2010.
13. M. Connors, P. Wiegert, and C. Veillet, "Earth's trojan asteroid", *Nature*, **475**, pp.481-483, 2011.

14. D.T. Britt, D. Yeomans, K. Housen, and G. Consolmagno, "An introduction to the dynamics of trojan asteroids", in *Asteroids III*, eds. W.F. Bottke Jr. *et al.*, U. of Arizona, Tucson, 2002.
15. J.G. Hills, "Computer simulations of encounters between massive black holes and binaries", *Astrophysical Journal*, **102**, pp.704-715, 1991.
16. M.L. Cosmo and E.C. Lorenzini (eds), "*Tethers In Space Handbook, Third Edition*", NASA, December 1997.
17. M. Finckenor, "Space tether", *Aerospace America*, **78**, 2005.
18. S. Bilen, "Space tethers", *Aerospace America*, **89**, 2007.
19. R.P.S. Oldenhuis, "Trajectory optimization for a mission to the solar bow shock and minor planets", Master's thesis, Astrodynamics & Satellite Systems, Aerospace Engineering, Delft University of Technology, January 2010.
20. R. Adams, "Two-burn escape maneuver for high Dv capable spacecraft", www.icarusinterstellar.org/papers/ASPW2010-2-Burn%28Adams%29.pdf. (Last Accessed 3 July 2013)

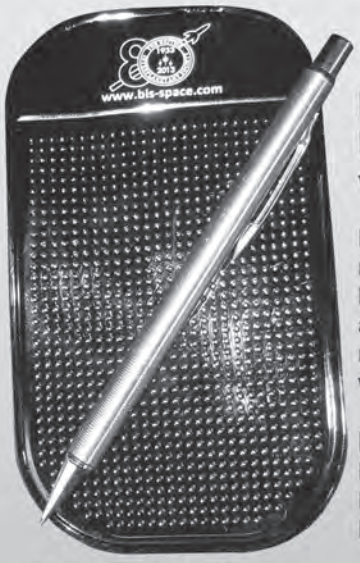
(Received 28 September 2011; Revision 16 March 2012 & 20 May 2013; Accepted 21 June 2013)

* * *

THE BRITISH INTERPLANETARY SOCIETY

FROM IMAGINATION TO REALITY

Summer 2013 Range



NEW

Non-slip Gel Pad (80th Anniversary edition)

Perfect for holding mobile phones, pens, keys, etc. Tactile adhesive that can be rejuvenated by a quick wash in warm water

Members £4.00
P&P: UK £1.50
EU £3.50
ROW £3.50

SKYLON Model

(Reaction Engines)
1:250 Scale, 33cm long.
Members £20.00
Non-Members £27.00
P&P: UK £3.50
EU £5.50
Rest of World £9.00

NEW

New Corporate BIS Tie (centre)

The new silk tie in maroon featuring the BIS comet insignia, the traveller across interplanetary space, is also used as the JBIS logo.

Comet Silk Tie (centre):
Members £15.00
Non-Members £20.00

BIS Silk Tie (right):
Members £15.00
Non-Members £20.00

BIS Polyester Tie (left):
Members £10.00
Non-Members £12.00

P&P: UK £2.50
EU £3.50
Rest of World £3.50

Sports Shirts

Polo shirts and T-shirts in navy or white with embroidered Society logo.

POLO SHIRTS:	T-SHIRTS:
Members £15.00	Members £10.00
Non-Members £18.00	Non-Members £12.00

Available in: XS, S, M, L, XL, 2XL, 3XL, 4XL, 5XL
P&P : UK £3.50 EU £5.00 Rest of World £7.50

80th Anniversary Mug

NEW

Featuring a cartoon strip by our very own President
Members: £5 P&P : UK £3.50 EU £5.50 ROW £9.00

Coming Soon....

Order Online Now - www.bis-space.com

Or by phone: BIS HQ office: +44 (0)20 7735 3160 Fax: +44 (0)20 7587 5118

BUILD A STARSHIP

STARSHIP CONGRESS

ANATOLE HILTON • DALLAS, TX

AUGUST 15-18, 2013

BIS Lectures and Meetings

Sampling the Mysteries of Space by Sample Return of Comets and Asteroids

Natalie Starkey

31 July 2013, 7 - 8.30 pm

Comets and asteroids act as time travel capsules, having captured and preserved the earliest Solar System material that formed 4.6 billion years ago. Here on Earth we are fortunate to receive free samples of comets and asteroids that arrive naturally as meteorites and dust particles at the Earth's surface. However, sampling the rocky, icy surfaces of comets and asteroids in space allows for the collected material to be understood in context, and provides pristine samples for analysis. In this talk I will discuss the many sample return space missions past, present and future that are pushing technological boundaries to return comet and asteroid samples to Earth. I will discuss the importance these missions have had on the science of the Solar System, and what new scientific developments and discoveries have been made.

68th Annual General Meeting

7 September 2013, 1 pm

The 68th Annual General Meeting of the Society will be held at the BIS HQ, 27/29 South Lambeth Road, London SW8 1SZ on Saturday 7 September 2012 at 1 pm followed by refreshments.

Admission to the AGM is open to Fellows only but all Members are welcome to join the discussion after the formalities of the AGM around 1.30 pm. Please advise in advance if you wish to attend.

JUpiter ICy moons Explorer (JUICE): The First Large ESA Cosmic Vision Mission

Athena Coustenis

12 September 2013, 7 - 8.30 pm

The Jupiter Icy Moons Explorer (JUICE) mission selected by ESA as the first large mission within the Cosmic Vision 2015-2025 plan, is being developed to address questions regarding the Jupiter system and its satellites, with a focus on the largest moon, Ganymede. The over-arching theme for JUICE is the emergence of habitable worlds around gas giants taking into account the requirements involving the presence of organic compounds, trace elements, water, energy sources and a relative stability of the environment over time.

For Europa, two targeted flybys are planned, with a focus on the chemistry essential to life, including organic molecules, and on understanding the formation of surface features and the composition of the non water-ice material, leading to the identification and characterisation of candidate sites for future in situ exploration.

The JUICE mission is planned to be launched in mid-2022, with a backup opportunity in August 2023. It will arrive at Jupiter in January 2030 after 7.6-years using an Earth-Venus-Earth-Earth gravity assist sequence and is foreseen to last for 3 and a half years.

From Imagination to Reality 2

14 September 2013, 9 am - 6 pm

An all-day Space Event.

Built on the BIS motto, From Imagination To Reality, looks at the ideas that first took hold in 'the imagination' – books, paintings, film and television, and sees how they have become 'the reality' – maybe not at the current time, but could very well happen in the future – even if it is the far future?

Based on the success of the first such event last year – From Imagination To Reality 2, examines the commercialisation of space, borrowing it's alternative title from a quote by that most famous of visionaries, and founder member of the BIS, Sir Arthur C. Clarke.

From Imagination To Reality 2 will include features on what certainly has come about, communications satellites; what is soon to come about, the first steps with space tourism; what should come about, the totally re-usable shuttle; what could potentially come about – mining the vast resources of the asteroids; what will eventually come about, 'Life on Mars'; and, perhaps the most grandiose of all, what if the fledgling ideas of space colonies or world-ships finally comes about? Could we – at some point – all be living in self-contained artificial worlds, on our way to the stars – as envisioned in so much fiction? From Imagination To Realty 2 examines the possibilities…space exploration and settlement. It will bring together scientists, policy makers, sociologists and political philosophers.

You can register online at www.bis-space.com or by requesting a registration form.

Man in the Solar System: A Very British Approach

Alan Bond

2 October 2013, 7.30 pm

Venue: Bath Royal Literary and Scientific Institute, Bath, BA1 2HN

In the 1980s a promising new propulsion concept was investigated in the UK leading to a 2-1/2 year intensive study of a single stage to orbit reusable space launch vehicle called HOTOL. HOTOL was studied by British Aerospace and Rolls-Royce but abandoned because of lack of Government support at the end of the 1980s.

The project was not abandoned however by its inventors and in 1989 Alan Bond, Richard Varvill and John Scott created Reaction Engines Ltd to continue to pursue the concept. The result has been the SKYLON Spaceplane and its SABRE powerplant. In studying the requirements for the Spaceplane a number of additional studies have been conducted into space transport infrastructure and missions to Mars. This talk will describe these studies and the experimental development which has taken place to date to bring SKYLON to fruition.

The Nervous System of a Starship

Pat Galea

17 October 2013, 7.30 pm

Venue: Bath Royal Literary and Scientific Institute, Bath, BA1 2HN

Sending a ship to the stars has many daunting technical challenges. Discussion often focuses on propulsion but control and communication over the vast distances and timescales involved is also a formidable challenge.

This lecture will discuss how advanced computation and communication technologies such as machine learning and gravitational lensing can be used to build ships which can work reliably and autonomously over centuries, and communicate their status and their data to Earth.

Readers are reminded that these Notices contain only a reduced description of the event. Full details can be found on the website at www.bis-space.com, where any updates are also carried.

Non-BIS Events

Starship Congress 2013

15-18 August 2013

Venue: Hilton Anatole, Dallas, Texas

www.icarusinterstellar.org/congress-announcement/

Lectures

Venue: Lectures will be held at BIS HQ, 27/29 South Lambeth Road, London, SW8 1SZ, unless otherwise stated.

Members can attend free of charge. Places must be booked in advance, online or by post. Each member may also obtain a free ticket for one guest subject to availability of space.

Non-Members are able to attend the Society's lectures for a fee. You can order a ticket online or by post (please make cheques payable to the British Interplanetary Society). If oversubscribed Society Members will be given priority.

If applying by post please send an sae. If applying via our website the confirmation receipt is your entry ticket.

If, for reasons outside its control, the Society is required to change the date or topic of a meeting, every effort will be made to avoid inconvenience to attendees either by notice of change in ***Spaceflight/JBIS***, on our website or by special advice to each participant.

JBIS

Journal of the British Interplanetary Society

jbis.org.uk