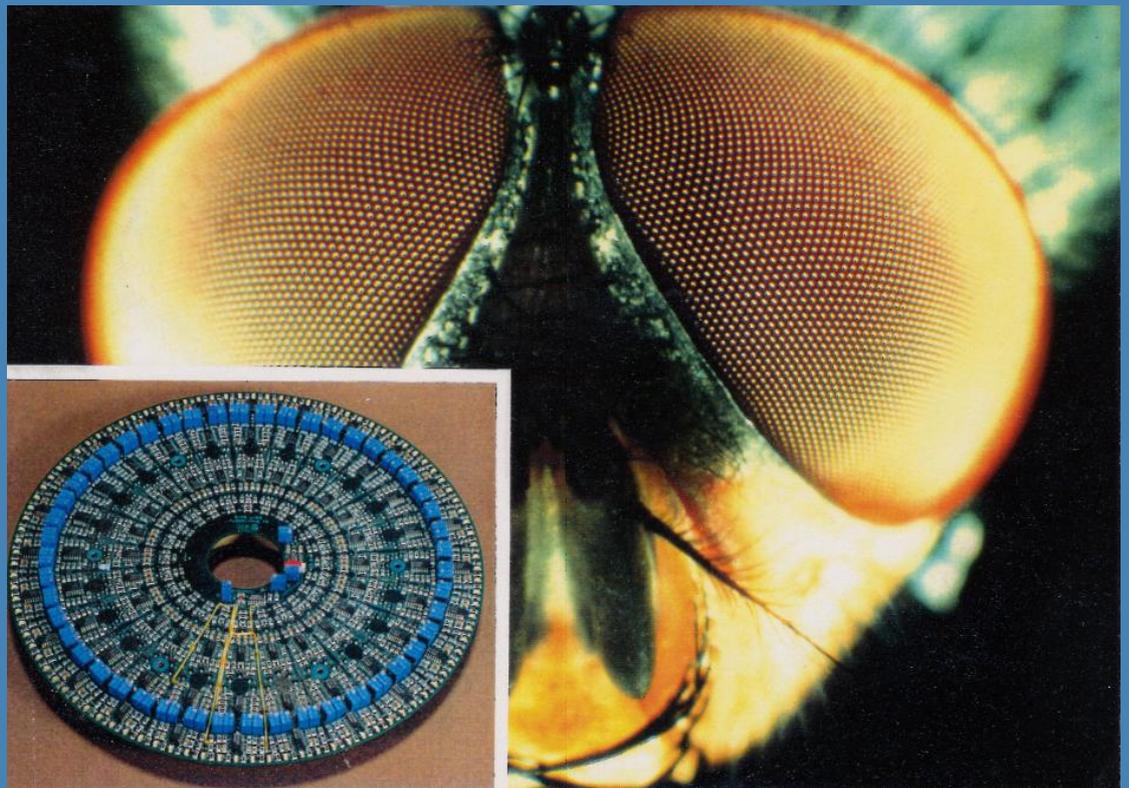


Acta Futura

Issue 3

Special Issue on Innovative System Concepts



Advanced Concepts Team

<http://www.esa.int/act>

Publication: Acta Futura, Issue 3rd (2009)

Editorial board: Jose M. Llorens Montolio
Dario Izzo
Christos Ampatzis
Pacôme Delva
Francesco Biscani

Published and distributed by: Advanced Concepts Team
ESTEC, DG-PI
2201 AZ, Noordwijk
The Netherlands
www.esa.int/gsp/ACT/publications/ActaFutura
Fax: +31 71 565 8018
D.O.I.: 10.2420/ACT-BOK-AF03

Copyright ©2009 - ESA Advanced Concepts Team

Typeset with X_YT_EX

Second Workshop on
Innovative Concepts

28-29 January 2008
Noordwijk, The Netherlands

**Organised by ESA
Advanced Concepts Team**



Organising Committee (ESA):

Claudio Bombardelli
Cristina Bramanti
Maria Johansson
François Nuyttens

Scientific Committee (ESA):

Luzi Bergamin
Daniela Girimonte
Dario Izzo
Jose M. Llorens Montolio
Naomi Murdoch
Marco del Rey Zapatero
Luca Rossini
Tobias Seidl
Leopold Summerer
Laura Torres Soto
Tamas Vinko

Workshop Agenda
ESTEC, 28-29 January 2008, Newton 2

28 January 2008

- 9:00 Welcome and Introduction
Leopold Summerer, ESA ACT
- 9:15 Keynote Address: "Technological Dynamics in World History"
Ian Inkster, Nottingham Trent University
-

PANEL 1 - Enabling Tomorrow's Pioneers

Chair: *Gianfranco Visentin, ESA*

Coordinator: *Cristina Bramanti, ESA*

- 9:45 Panel Introduction
- 9:50 Visual guidance based on optic flow
Nicolas Franceschini, Université de la Méditerranée, Marseille, France
- 10:15 Spin-ins for a new propulsion era?
Steve B. Gabriel, University of Southampton, UK
- (Coffee break)
- 11:00 PEIS concepts integrating robots into smart environments
Alessandro Saffiotti, University of Örebro, Sweden
- 11:25 Integration of cellular biological structures into robotic systems
Klaus-Peter Zauner, University of Southampton, UK
- 11:50 Panel Discussion

(Lunch break)

PANEL 2 - "Vis Viva" from Space

Chair: *Carla Signorini, ESA*

Coordinator: *Claudio Bombardelli, ESA*

- 13:50 Panel Introduction
- 13:55 Extracting power out of an active environment (e.g. vibrations)
Steve Beeby, University of Southampton, UK
- 14:20 Biomass fuel cells
Aarne Halme, University of Helsinki, Finland
- 14:45 Orbital energy of natural satellites converted into permanent power for spacecraft
Jesús Pelaez, Universidad Politecnica de Madrid, Spain
- 15:10 Harnessing power from solar wind particles captured in the Van Allen belts
Andrei Yakovlev, St. Petersburg State University, Russia
- 15:35 Panel Discussion

PANEL 3 - "Driving Evolution"

Chair: *Leopold Summerer*, ESA

Coordinator: *François Nuyttens*, ESA

- 16:20 Panel Introduction
- 16:25 Mechanochemistry: targeted delivery of single molecules
Ann-Sophie Duwez, University of Liège, Belgium
- 16:50 Bio-inspired surface-interaction (dry attachment, frictional sliding)
and locomotion systems
Stanislav Gorb, Max Planck Institute for Metals Research,
Stuttgart, Germany
- 17:15 Carbon Nanotubes: future horizons
Siegmar Roth, Max Planck Institute for Condensed Matter Physics,
Stuttgart, Germany
- 17:40 Panel Discussion

(Visit to ESTEC test facilities 18:10 - 18:45)

29 January 2008

PANEL 4 - Discovering Natural Paths

Chair: *Bernard Foing*, ESA

Coordinator: *Dario Izzo*, ESA

- 9:45 Panel Introduction
- 9:50 Electrostatic swarm control
Lorenzo Pettazzi, ZARM - University of Bremen, Germany
- 10:15 Invariant satellite motions
Marco Sabatini, University of Rome, Italy
- (Coffee break)
- 11:00 Pulsar navigation
Josep Sala, Technical University of Catalonia, Spain
- 11:25 Insect navigation and path finding
Tobias Seidl, ESA Advanced Concepts Team, The Netherlands
- 11:50 Panel Discussion
- (Lunch break)
-

CLOSING PANEL - Inputs to "Space 2030"

Chairs: *Alain Dupas*, Prospective 2100; *Géraldine Naja-Corbin*, ESA

Coordinator: *Leopold Summerer*, ESA

- 14:00 Panel Discussions
- 15:30 Workshop wrap-up by organising committee

Contents

Foreword	11
Panel 1: “Creating Tomorrow’s Pioneers”	13
Towards automatic visual guidance of aerospace vehicles: from insects to robots <i>Nicolas Franceschini</i>	15
The Concept of PEIS-Ecology: Integrating Robots in Smart Environments <i>Alessandro Saffiotti</i>	35
Integration of Cellular Biological Structures Into Robotic Systems <i>J. Gough, G. Jones, C. Lovell, P. Macey, H. Morgan, F. D. Revilla, R. Spanton, S. Tsuda, K.-P. Zauner</i>	43
Panel 2: “ <i>Vis Viva</i> ” from Space	51
Kinetic Energy Harvesting <i>S. P. Beeby, R. N. Torah, M. J. Tudor</i>	53
Biomass Based Fuel Cells - Application to Manned Space Exploration <i>A. Halme</i>	63
Orbital energy of natural satellites converted into permanent power for spacecraft <i>J. Peláez</i>	71
Harnessing of the power of the solar wind particles captured in the Van Allen belts <i>E. K. Kolesnikov, A. B. Yakovlev</i>	81
Panel 3: “Driving Evolution” – A paradigm shift in the balance between nature and technique? Consequences and opportunities for space?	89
Playing with forces and interactions to manipulate single molecules <i>Bouzid Menaou, Anne-Sophie Duwez</i>	91
Panel 4: “Discovering Natural Paths”	97
Electrostatic Force for Swarm Navigation and Reconfiguration <i>Lorenzo Pettazzi, Hans Krüger, Stephan Theil, Dario Izzo</i>	99
Invariant Relative Satellite Motion <i>M. Sabatini, D. Izzo, G. Palmerini</i>	107
Pulsar Navigation <i>Josep Sala, Andreu Urruela, Xavier Villares, Jordi Romeu, Sebastià Blanch, Robert Estalella, Josep M. Paredes</i>	115
Insect navigation and path finding <i>Tobias Seidl</i>	125
Workshop Conclusions	131

Foreword

Two years have passed since the ESA Advanced Concepts Team (ACT) has organised the 1st Workshop on Innovative System Concepts (proceedings published as a special issue of this journal). Within these two years, the ACT has completely renewed its team of researchers and, thus, many of its ideas and concepts. In line with the spirit of pushing forward the limits of current concepts, methods and technologies for space, the team is glad to present with this publication the outcome of its 2nd Workshop on Innovative Concepts, held on the 28 and 29 of January 2008 at ESTEC in the Netherlands.

The workshop presented new ideas and concepts explored by the team, advances in science that were not yet part of the core space domain and topics considered as potentially interesting for future space activities to a larger audience of essentially space scientists and engineers. The contributions to the workshop in form of discussion fora, presentations and scientific papers were organised around four meta-themes: *i*) Creating tomorrow's Pioneers, *ii*) *Vis viva* from space, *iii*) Driving evolution and *iv*) Discovering Natural Paths. These themes reflect on the one hand the thematic orientation of the team's research interests and demonstrate the cross-disciplinary approach for its research: None of them can be classified according to the conventional research lines present in academia. Indeed, they can only exist and flourish inside a multidisciplinary team where boundaries are crossed and disciplines broken down.

Reflecting the same spirit, the rationale behind each of the meta-themes has been reproduced in the present issue. Each of them sets up a chapter embracing a collection of papers written by the speakers of the workshop. Unfortunately, this issue cannot reflect the lively and fruitful discussions that took place during the different panel discussions of the workshop, but contains the transcript of the closing remarks of Ms. Géraldine Naja-Corbin, Head of Institutional Affairs and Strategic Studies at ESA. More important than the trace of the discussions however, we hope to see their impact on future space activities.

This issue of *Acta Futura* presents a snapshot in time and a selection of limited research subjects by and with the team. More publications and studies are available from the ACT website.

Advanced Concepts Team

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

Panel 1: “Creating Tomorrow’s Pioneers”

THE need for “Discovery” and the drive to “Explore” have always been fundamental to characterizing the development of the human race. There is a clear relationship between our technological development and our abilities to explore and make discoveries in ever more remote locations. The two are intertwined - each in turn, furthering the other. The Renaissance in Europe spawned technological advances which led to an “Age of Exploration”. From early 15th century onwards European ships and their navigators were so technologically advanced as to be capable of reaching anywhere but the most extreme regions of the globe. By the early 17th century their range of exploration was limited only by their technologies. Advancement into the less hospitable regions of the world had to wait until the supporting technology was available.

In the 20th century, Space Exploration has been a new human quest, with the goal to discover the nature of the universe beyond Earth. Our pioneers have been pushing slowly forward, on limited resources, expanding our boundaries out into space. Perhaps we sit now on the cusp of a new “Age of Exploration”. Today, early in the 21st century, there is an indication that we are at a point where the combinations of an array of new technologies leave us poised to send out new waves of supremely equipped expeditions. Our new pioneers will have unprecedented technological advantages: where will their limit be? The extension of current technology might not be enough for enabling future exploration challenges, and will reach a natural limit if new capabilities, inspired by emerging discoveries and disciplines, are not developed. The synergy between different fields such as power, computing, nanotechnology, biotechnology, communications, networking, robotics, materials might help to drive the progress. Meta-capabilities, as a result of this synergy, can then be exploited in building the Pioneer space vehicles of tomorrow for a new wave of discovery.

This panel will address how new enabling capabilities can be derived from the cross-fertilisation between different disciplines such as biology, robotics, mechanics, chemistry, artificial intelligence, material science. Will tomorrow’s robotic pioneers use bio-inspired electronics to find their path and react to the unforeseen? Will they use innovative visual sensors inspired by insects,

which have capabilities far exceeding anything that we can currently build in many areas? Are we going to use robots capable to genuinely and productively interact with the humans and their environment? Propulsion is at the heart of all space missions, but propulsion technology has not seen major breakthroughs since its early days - could new inspiration for propulsion derived from terrestrial technologies provide for the extra “kick” that would enable so many more space applications?

New sources of inspiration seems to be needed to boost innovative ideas leading to new fields of research and cooperation between scientists. The goal of this panel is to discuss the different research concepts with an emphasis on their potentials and applicability in a mid to long-term time scale.

Cristina Bramanti
Panel coordinator



Acta Futura 3 (2009) 15-34

DOI: 10.2420/AF03.2008.12

**Acta
Futura**

Towards automatic visual guidance of aerospace vehicles: from insects to robots

NICOLAS FRANCESCHINI*

*Biorobotics, Movement Science Institute (CNRS) and University of the Mediterranean,
163 Avenue de Luminy (CP938) Marseille F-13288, cedex 9, France*

Abstract. Equipped with a less-than-one-milligram brain, insects fly autonomously in complex environments without resorting to any Radars, Ladars, Sonars or GPS. The knowledge gained during the last decades on insects' sensory-motor abilities and the neuronal substrates involved has provided us with a rich source of inspiration for designing tomorrow's self-guided vehicles and micro-vehicles, which are to cope with unforeseen events on the ground, in the air, under water or in space. Insects have been in the business of sensory-motor integration for more than 100 millions years. They can teach us useful tricks for designing agile autonomous vehicles at various scales. Constructing a "biorobot" first requires exactly formulating the signal processing principles at work in the animal. It gives us, in return, a unique opportunity of checking the soundness and robustness of those principles by bringing them face to face with the real physical world. Here we describe some of the visually-guided terrestrial and aerial robots we have developed on the basis of our biological findings. All these robots react to the *optic flow* (i.e., the angular speed of the retinal image). Optic flow is sensed onboard the robots by miniature vision sensors called Elementary Motion Detectors (EMDs). The principle of these electro-optical velocity sensors was derived from opti-

cal/electrophysiological studies where we recorded the responses of single neurons to optical stimulation of single photoreceptor cells in a model visual system: the fly's compound eye. Optic flow based sensors rely solely on contrast provided by reflected (or scattered) sunlight from any kind of celestial bodies in a given spectral range. These passive sensors and systems offer potential applications to manned or unmanned spacecraft, from robotic landers and rovers to asteroid explorers and satellite docking, with interesting prospects in weight-reduction and low consumption.

1 Introduction

Animals and humans are natural «vehicles» [13], able to move about autonomously in complex environments. Insects, in particular, provide evidence that physical solutions to elusive problems such as those involved in robots' visually-guided locomotion existed millions of years before roboticists started tackling these problems in the 20th century. Over the past two decades, some research scientists have been attempting to tap insect biology for ideas as to how to design smart visually-guided vehicles, e.g. [13, 63, 103, 86, 84, 43, 25, 44, 7, 116, 20, 85, 88, 90, 89, 34, 122, 92, 2, 87, 83, 96, 66, 82, 16, 130, 19, 55, 68, 67, 97, 138, 4, 94, 111, 95, 139, 110, 52, 144, 113, 70, 119]. Insects have quite a wide behavioral

*E-mail address: nicolas.franceschini@univmed.fr

repertoire and can teach us how to cope with unpredictable environments using smart sensors and limited processing resources. Flying insects, in particular, often attain a level of agility that greatly outperforms that of both vertebrate animals and present day mobile robots. Insects' sensory-motor control systems are masterpieces of integrated optonics, neuronics and micromechatronics. Their neural circuits are highly complex – commensurate with the sophisticated behaviour they mediate – but unlike most (no less complex) vertebrate neural circuits, they can be investigated at the level of single, *uniquely identifiable neurons*, i.e., neurons that can be reliably identified in all the individuals of the species on the basis of their location in the ganglion, their exact shape and their consistent electrical responses, review: [124, 65, 99, 18, 24]. This great advantage of insect versus vertebrate neuroscience enables insect neuroscientists to accumulate knowledge during anything from a few days to several decades about a given individual neuron or a well defined neural circuit. This explains why many of the robots that emulate part of a nervous system were inspired by arthropods and particularly insects [7, 16, 122, 82, 19, 138, 139, 4, 8, 52]. The biologically based robots that we have been constructing since 1985 have largely contributed to creating the field of Biorobotics, in which natural principles or systems are *abstracted* from an animal and implemented in the form of a hardware physical model. After an introduction to the fly visual system (Section 2), we will be briefly described some of the insect based robots we conceived and constructed (Sections 3 and 4). Beyond their possible usefulness as intelligent machines, these physical models make it possible to subject biological hypotheses to the rigorous tests imposed by the real world [136, 137, 46]. Section 5 will deal with some of the “biological returns” that came out from this biorobotic approach. Section 6 will deal with possible applications of these findings to aerospace.

2 Fly visual microcircuits

Flies are agile seeing creatures that navigate swiftly through the most unpredictable environments, avoiding all obstacles with little conventional aerospace avionics. Equipped with “only” about one million neurons and “only” 3000 pixels in each eye, the housefly, for example, achieves e.g., 3D navigation and obstacle avoidance at an impressive 700 body-lengths per second. All this is achieved, surprisingly, without any connections of the

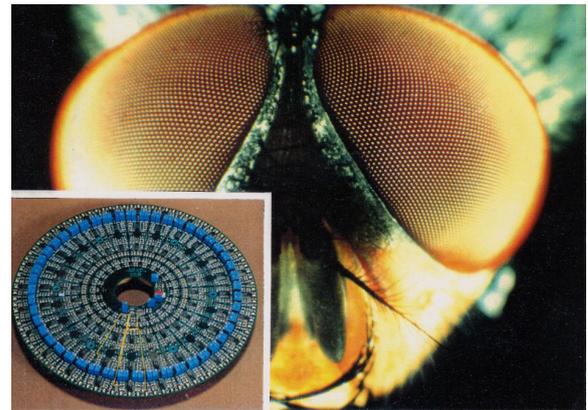


FIGURE 1. Head of the blowfly *Calliphora erythrocephala* (male) showing the two panoramic compound eyes with their faceted cornea. Each facet lens is the front-end of an ommatidium that contains a small group of photoreceptor cells (Fig. 2). There are as many sampling directions (pixels) in each eye as there are facets. This photograph was taken with a laboratory made Lieberkühn microscope based on two parabolic mirrors extracted from bicycle lights. Inset: photograph of the retinotopic circuit used by the Robot-Fly (Fig. 5a) to guide itself at a high speed towards a target while avoiding obstacles on its way. The routing of this circuit is shown in Fig. 7 extracted from [35].

animal to a super-computer and an external power supply. The impressive lightness of the processing system at work onboard a fly or a bee makes any roboticist turn pale once he/she realizes that these creatures achieve many of the behaviours that have been sought for in the field of autonomous robotics for the last 50 years: dynamic stabilization, 3D collision avoidance, tracking, docking, autonomous landing, etc.

The front end of the fly visual system consists of a mosaic of facet lenslets (Fig. 1) and an underlying layer of photoreceptor cells forming the retina proper (Fig. 2). Insects' photoreceptor cells, once dark-adapted, are known to respond to single absorbed photons by a miniature potential («bump») of a few millivolts amplitude. They are more sensitive and reliable than any photomultiplier ever built. Flies possess one of the most complex and best organized retinas in the animal kingdom. It has been described with unprecedented details, with its six different spectral types of photoreceptor cells, polarization sensitive cells, and sexually dimorphic cells. The analysis has revealed a typical *division of labour* within each ommatidium:

- The two central photoreceptor cells, R7 and R8,

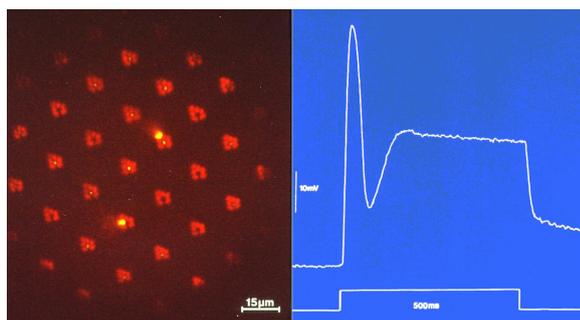


FIGURE 2. (a) Photoreceptor cells of the housefly retina observed *in vivo* using the “corneal neutralization” technique [39, 29, 30] combined with epi-fluorescence microscopy under blue excitation [41]. Each cluster of 7 spots stands for the seven rhabdomeres (dark spots in the inset of Fig. 4a, right) located beneath each facet. While the peripheral rhabdomeres (R1–R6) all fluoresce red, the central rhabdomere (R7) falls in three classes: green fluorescent (yR7), nonfluorescent (pR7) and red fluorescent (rR7). The latter type was shown to occur exclusively in the dorsal part of the male eye [38]. The fluorescence colors are linked to the specific visual pigment present in each rhabdomere and constitute a genuine “color code” to their spectral sensitivities [41, 38, 31] (from [31]). (b) We recorded the receptor potential delivered by an individual photoreceptor cell (cf [79]). Using ultrafine micropipettes filled with a fluorescent dye, we injected the dye into the cell immediately after measuring its spectral sensitivity $S(\lambda)$. The injected cells (here R2 and R3, see the numbering of cells in Fig. 4a, inset) were recovered *in vivo*, allowing fast and unequivocal assignment of a given $S(\lambda)$ to a given cell type (from [31])

display specific spectral sensitivities [88] and are therefore thought to mediate color vision [54, 31, 53]. The ommatidia are spectrally *heterogeneous* as regards R7 and R8, which can be mapped out accurately *in vivo* using the “corneal neutralization” technique [39, 41, 38] (see Fig. 2a). 70% of the R7 cells are of the « yellow » (y) type (they appear green in Fig. 2a) and 30% of them are of the « pale » (p) type (they appear dark in Fig. 2a) [74, 41, 38]). To each pR7 and yR7 receptor corresponds a specific pR8 and yR8 tandem companion, both companions showing yet other spectral sensitivities [54]. This complex but analyzable pattern actually provided the first evidence that color vision in insects relies on *spectrally heterogeneous ommatidia* [31, 32, 53], a finding that was corroborated only recently for insects with fused rhabdoms, such as bees and butterflies [105, 15, 134]. Flies achieve specific receptor spectral sensitivities through the

expression of different opsin genes in photoreceptors [102].

- **The outer 6 receptor cells (R1–R6)** contribute in particular to *motion detection* [73, 109, 17, 59]. Consistent with this function where contrast sensitivity is at a premium, the R1–R6 photoreceptors make for a high sensitivity (“scotopic”) system whose enhanced signal-to-noise ratio is due, in part to the specific neural wiring (the “*neural superposition principle*”, [12, 72, 75]) and also to their panchromatic, ultraviolet-enhanced spectral sensitivity [76]. As inputs to the motion detection pathway, R1–R6 receptors, as we will see, play a major role in insect visual guidance based on “*optic flow*” (OF).

Flying insects avoid collisions with obstacles and guide themselves through their complex surroundings by processing the *optic flow* (OF) that is generated on their eyes as a consequence of self-motion. In the animal’s reference frame, the OF is the *angular speed* ω (expressed in rad/s) at which each contrasting object of the environment moves past the animal [49, 80, 77]. Fig. 3a illustrates the case for an insect flying in translation over a terrain. Current evidence shows that the insect’s nervous system is able to perform the complex task of extracting the information necessary for short range navigation from the optical flow field [23, 21, 51, 140, 132, 17, 81, 71, 22, 118, 28] and to transmit this information to the thoracic wing muscles in a « fly-by-wire » mode. This ability results from the insect head being equipped with smart sensors called *motion detecting neurons*, which are able to gauge the *relative motion* between the animal and environmental features [56, 59, 135, 11].

The fly is one of the best animal models currently available for studies on motion perception [107, 50, 73, 17, 109, 32, 58, 45, 123, 78, 28]. A great deal has already been learned from neuroanatomical and neurophysiological studies on the 3rd optic ganglion, part of which (the *lobula plate*) appears as a genuine “visual motion processing center”. This region comprises approximately 60 *uniquely identifiable* neurons dedicated in particular to: (i) analysing the movement of the retinal image, i.e., the optic flow field that results from the animal’s walking or flying, and (ii) transmitting the result of this analysis via descending neurons to the thoracic interneurons that will ultimately drive the wing-, leg-, and head-muscles [124, 123, 56, 57, 58, 78, 11, 28]. The *lobula plate tangential cells* are large-field collator neurons that pool the input signals from many retinotopic

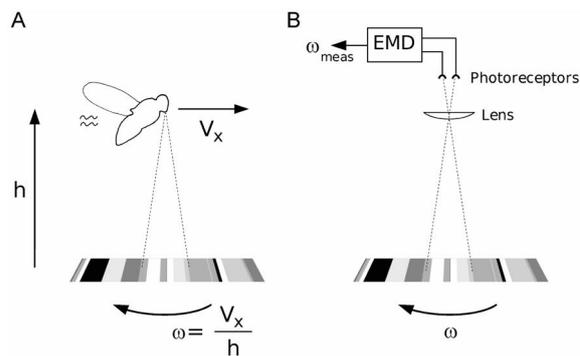


FIGURE 3. Definition (ω) and measurement (ω_{meas}) of the ventral optic flow (OF) perceived by an animal or a robot flying in translation in the longitudinal plane. (a) The ventral optic flow perceived by an animal flying at groundspeed V_x and groundheight h is the angular speed ω at which a point in the ground texture directly below seems to move in the opposite direction. By definition, ω (in rad/s) is the ratio between groundspeed and groundheight. The one-dimensional randomly textured ground shown here is a magnified sample of that shown below Fig. 9a. (b) The OF sensor used onboard the robot OCTAVE (Fig. 5b) comprises a microlens and two photoreceptors driving a fly-based Elementary Motion Detector (EMD). The output ω_{meas} from the OF sensor serves as a feedback signal in the control scheme shown in Fig. 8 (from [46]).

“Elementary Motion Detectors” (EMDs) [56, 57]. Although the detailed neural circuitry underlying *directionally selective motion sensitivity* still remains elusive in all animals – vertebrates and invertebrates – two types of smallfield columnar neurons have been identified in flies, the transmedullary neuron Tm1 and the bushy lobula neuron T5, which seem to be major players for conveying smallfield motion information down to the large-field lobula plate neurons [27, 62]. Regardless of the detailed EMD neuronal circuitry, the problem we addressed in the 1980’s was the *functional principle* underlying an EMD. Taking advantage of the micro-optical techniques we had developed for analysing the fly retina at the single photoreceptor level [39, 40, 29, 30, 31], we were able to stimulate a single EMD in the eye of the living insect by applying optical stimuli to single *identified* photoreceptor cells on the retinal mosaic (Fig. 4), while recording from an *identified* motion sensitive neuron, called H1, in the lobula plate of the housefly [79, 25]. We applied pinpoint stimulation to two neighboring photoreceptors (diameter 1 μm) of a single ommatidium by means of a special stimulation instrument (a hybrid between a microscope and a telescope, see Fig. 4, left), in which the main objective lens was the facet lens itself

(diameter 25 μm , focal length 50 μm). This optical instrument [45] served (i) first to select a given facet lens (Fig. 4a), (ii) to select two of the seven receptors (namely R1 and R6) and illuminate them *successively* with 1 μm -light spots. This *sequential microstimulation* produced an “apparent motion” simulating a real motion within the small visual field of an ommatidium. Although, in this type of experiment, only two out of the 48,000 photoreceptor cells of the visual system received light, the H1-neuron responded by a conspicuous *increase* in spike rate, as long as the phase relationship between the two stimuli mimicked a movement occurring in the *preferred* direction (see Fig. 4c). When the sequence mimicked a movement in the opposite, (“null”) direction, H1 showed a marked *decrease* in its resting discharge or did not respond at all [109, 32]. H1 did not respond either when the same sequence was presented to a pair of receptors (such as R1 and R2, or R1 and R3) aligned *vertically* in the eye (consistent with the fact that H1 is not sensitive to vertical motion [56, 58]), or when one of the two selected photoreceptors was the central cell R7 (confirming that this cell does not participate in motion detection [73, 109, 59]).

From many experiments of this kind, in which carefully planned sequences of light steps and/or pulses were applied to the two receptors, we established the EMD *block diagram* and characterized each block’s dynamics and nonlinearity [32, 45, 33]. The scheme we arrived at departs markedly from the popular Hassenstein-Reichardt *correlation* model that was derived 50 years ago from *behavioural* experiments carried out on walking or flying insects [106, 107, 17, 108]. While not unveiling the details of the EMD neural circuit, our analysis at the single cell level allowed the EMD principle to be understood, paving the way for its transcription into another, man-made technology. It is on this basis that, in the mid 1980’s, we designed a miniature electronic EMD whose signal processing scheme approximated that of the biological EMD [36, 9]. Specifically, our scheme relies on measuring the time of travel of a contrasting feature between two neighboring pixels. This time of travel is further processed to give an output that grows as a monotonic function of the angular velocity, with little influence of the pattern contrast and spatial frequency. Our scheme is not a *correlator scheme* and rather corresponds to the class of *token-matching schemes*, as defined by S. Ullman’s [126]. It is therefore a true *angular velocity sensor* [36] that can serve as a worthy OF sensor (see Fig. 3b). Over the years, we miniaturized this EMD using a variety of technolo-

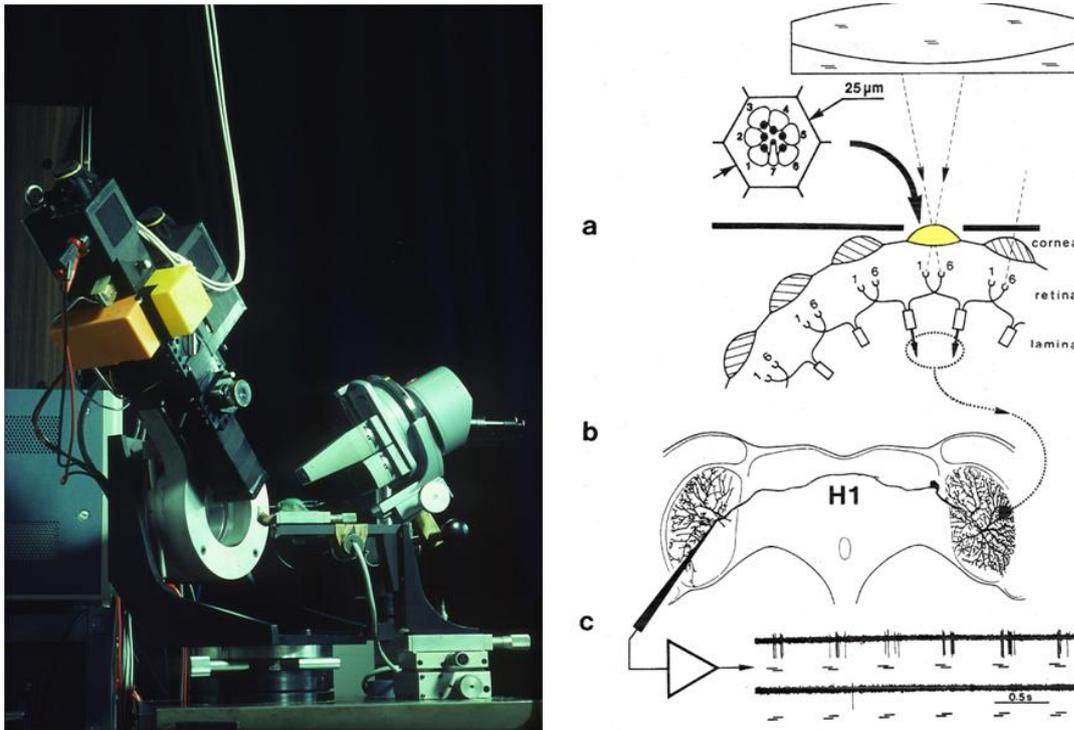


FIGURE 4. (left): Triple-beam incident light “microscope-telescope” built to analyze motion detection at the most elementary level. The instrument successively delivers a $1\ \mu\text{m}$ -light spot to two neighboring photoreceptor cells, R1 and R6, within a single ommatidium (see Fig. right, (a) inset). A microelectrode records the electrical response of the motion sensitive neuron H1 to this “apparent motion” [109, 32, 45] (right): Information flow from the objective of the “microscope-telescope” down to the H1 neuron, whose spiking activity is recorded contralaterally. Presenting two flashes (each with a duration of 100 ms) successively to the pair of receptors R1 and R6 consistently induced a prominent spike discharge in H1 whenever the sequence mimicked motion in the preferred direction (c, top trace) and no response when the opposite sequence was presented (c, bottom trace) (from [45]).

gies [36, 44, 115, 104, 3] and all the robots described below (Fig. 5) were equipped with either version of these fly-derived OF sensors (Fig. 6, top). A very similar EMD principle has been proposed independently a decade later by C. Koch’s group at CALTECH, who spread it under the name “facilitate and sample” velocity sensor [69]. These authors patented a smart analog VLSI chip based on this principle – without any reference to a possible inspiration from the fly [117].

3 Fly-inspired, visually-guided terrestrial robots

In the mid 1980’s, we started designing a robot to demonstrate how an agent could possibly navigate in a complex environment on the basis of optic flow. The robot was equipped with a planar compound eye and

a fly-inspired EMD array [103]. The latter was used to sense the OF generated by the robot’s own locomotion among stationary objects. The 50-cm high “robot-mouche” (*Robot-Fly* in English) that we realized in 1991 (Fig. 5a) was the first OF-based, completely autonomous robot able to avoid contrasting obstacles encountered on its way, while traveling to its target at a relatively high speed (50 cm/s) [10, 43, 44, 42]. The Robot-Fly was also based on ethological findings on real flies, whose most common flight trajectories were shown to consist of straight flight sequences interspersed with rapid turns termed *saccades* [23, 59, 133, 118, 125]. Straight flight sequences performed at speed V near an obstacle located at distance D and azimuth ϕ with respect to the heading direction generate in the robot’s eye a *translational* optic flow ω expressed as follows [142]:

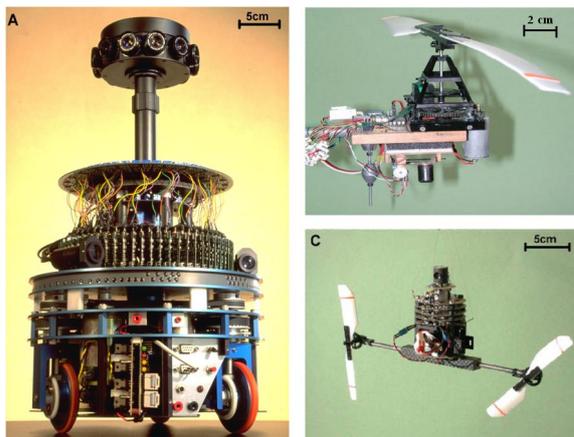


FIGURE 5. Three of the visually-guided robots designed and constructed at the Laboratory on the basis of our biological findings on visuomotor control in flies. (a) The Robot-Fly incorporates the compound eye shown in Fig. 6 for obstacle avoidance, and a dorsal eye for detecting the light source serving as a goal. This robot (height: 50 cm; weight 12 kg) reacts to the optic flow generated by its own locomotion amongst obstacles [10, 43, 44, 42]. It is fully autonomous as regards its processing and power resources. (b) The robot OCTAVE (Optic flow-based Control system for Aerospace Vehicles) is a 100-gram rotorcraft equipped with a 2-pixel ventral eye sensing the OF on the terrain below (see Fig. 3b, 8, 9). This self-sustained aerial creature is tethered to a light rotating arm that allows only three degrees of freedom: forward and upward motion and pitch. The robot lifts itself and circles around a central pole at speeds up to 3 m/s. It ascends or descends depending on the ventral optic flow it measures (from [111, 112, 113]). (c) The robot OSCAR (Optic flow based Scanning sensor for the Control of Aerial Robots) is a 100-gram, twin-engine aircraft equipped with a two-pixel frontal visual system that relies on visual motion detection and on a microscanning process inspired by the fly [37]. It is tethered to a 2-meter-long nylon wire secured to the ceiling of the laboratory. Vision and rate gyro signals combine onboard to make OSCAR fixate and track a target (a dark edge or a bar) with hyperacuity at speeds of up to 30°/s (from [128]).

$$\omega = V \sin(\phi)/D \quad (1)$$

This means that an animal (or a robot) able to measure ω could recover the distance D to the obstacle, provided it would know its own speed V . The robot Fly proceeds by performing a sequence of *purely translational* steps ΔL (length 10 cm; duration 200 ms) at a speed set at $V=50$ cm/s (via the wheel encoders). By the end of each step, the panoramic EMD-array has drawn up a map of the local obstacles, which is expressed in polar coordinates in the robot's reference frame. The next course

to be steered is immediately given, generating an “eye + body saccade” in the new direction. Vision is inhibited during each saccade by a process akin to “saccadic suppression”, a process that has also been documented in an insect [143]. The robot was tested wandering about autonomously in an arena in which obstacles (vertical poles) were arranged at random. Its jerky trajectory is very reminiscent of the flight trajectories recorded on real flies [133, 118, 125, 26]. The Robot-Fly is able to adapt to novel or changing environments without any need for maps or learning phases of any kind. The robot must move in order to be able to see. No “path planning” is required: the only thing that is “planned” is the direction of the next step, on the basis of what was sensed during the previous step and taking into account the direction of the target.

The Robot-Fly actually views the world through a horizontal ring of facets (Fig. 6c). Any two neighboring facets drive an EMD, and a total of 114 EMDs analyze the (self-generated) optic flow in the azimuthal plane. One peculiarity of the robot is that its compound eye is endowed with a *resolution gradient* such that the interommatidial angle $\Delta\phi$ increases according to the sine law of Eq. (1) as a function of the eccentricity ϕ [103, 44]. Once embedded in the anatomical structure of the eye, this resolution gradient “compensates” for the sine law inherent to the translational OF (Eq. (1)) and ensures that any contrasting feature will be detected if, during a robot's translation by ΔL , it enters the robot's “circle of vision”, the radius R_v of which increases linearly with ΔL :

$$R_v = k \Delta L \quad (2)$$

A similar nonuniform sampling is a key feature of many natural visual systems, including the human and fly visual systems. Sampling the environment *nonuniformly* in this way makes it possible to design the underlying EMD array *uniformly*, each EMD having the same time constants as its neighbors [44]. The Robot-Fly operates on the basis of a brainlike, parallel and analog mode of signal processing. Fig. 7 shows the odd routing pattern that connects the thousands of analog devices that blend together the input signals from the compound eye and those arising from the dorsal, target-seeking eye, eventually delivering a single analog output: the steering angle.

Simulation studies have shown that a robot of this kind can automatically adjust its speed to the density of the obstacles present in the environment if, instead of

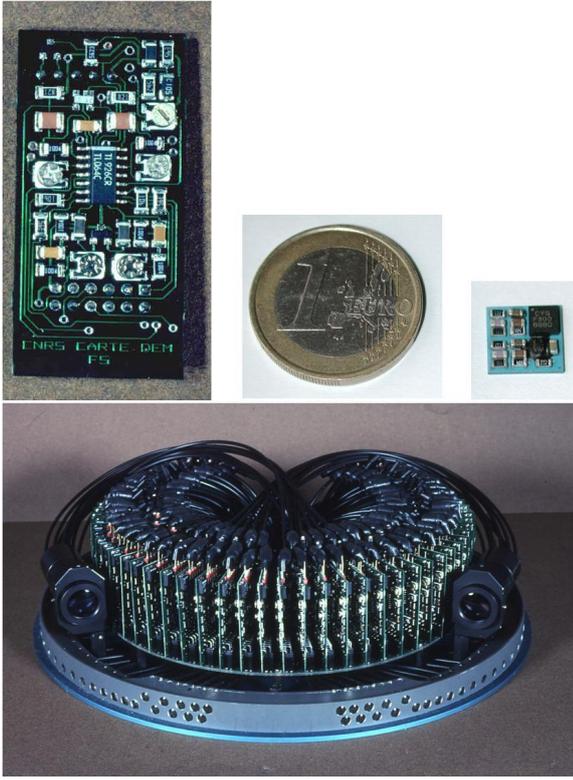


FIGURE 6. (top) Opto-electronic Elementary Motion Detector (EMD) derived from the functional diagram of a fly EMD established via physiological analysis (Fig. 4). The left, purely analog version of the EMD dates back to 1989 and weighs 5 grams. Our current EMDs are smaller hybrid (analog+digital) versions. The smallest one (top right), built in LTCC (low temperature co-fired ceramics) is only 7 mm x 7 mm and weighs only 0,2 grams (from [104]). Others are based on FPGAs [3]. (bottom) Compound eye with its faceted array and its 114 EMDs (of the type shown at the top left). Despite its small number of pixels, this artificial vision system is able to drive-by-wire the Robot-Fly (Fig. 5a) at high speed (50 cm/s), in a fully autonomous mode, through a field of unpredictable obstacles. Notice the resolution gradient from the front to the periphery of the visual field, which makes the sine function in Eq. (1) "embedded in the eye structure" (from [44]).

imposing on the robot constant ΔL translation steps (as described above), one imposes constant Δt translation times [85]. During any one Δt , the robot will then cover a distance ΔL proportional to its current speed V :

$$\Delta L = V \Delta t \quad (3)$$

From Eq. (2) and (3), one obtains:

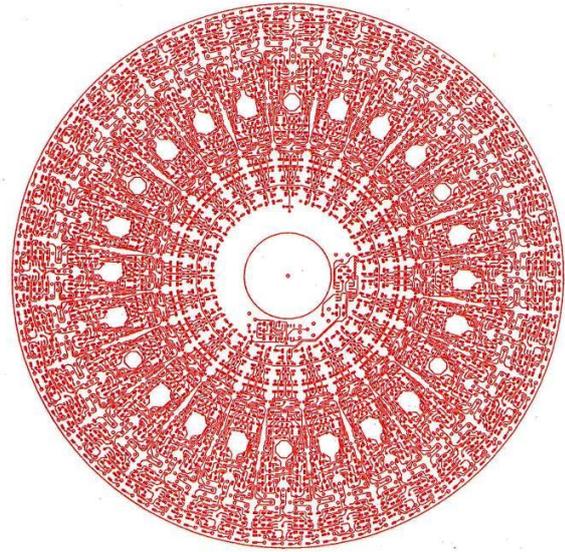


FIGURE 7. Routing diagram of one face of the printed circuit board (PCB) that integrates information about the obstacles and the target on-board the "robot-mouche" (Fig. 5a). This six-layered PCB has 210 parallel inputs (114 EMD inputs + 96 inputs from the target seeker) and a single output (near the centre of the pattern). The latter gives (in Volts) the next steering angle required to reach the target while avoiding all obstacles. The mosaic layout of this circuit is reminiscent of the neural architecture of insects' visual ganglia: the rose-window-like pattern results from the retinotopic projection of numerous repeat units (from [44]).

$$R_v = k V \quad (4)$$

Eq. (4) means that the robot will now be able to detect (and therefore avoid) obstacles within a range R_v that increases proportionally to its current speed V . The simulated robot was shown, making a detour around a dense forest, automatically accelerating in a clearing and automatically braking before traversing another, less dense forest that it "cautiously" entered at a reduced speed [85].

4 Insect inspired visually guided aerial robots

A similar motion detection principle can be used to guide a flying agent, have it follow a rough terrain and land [88, 96]. The latter paper [96] was something of a breakthrough as it showed for the first time how it is possible to follow a terrain and land *without knowledge*

of *groundspeed* and *groundheight*, in much the same way as the honeybee seems to operate [121, 120]. Following simulation experiments, the principle was first validated on-board FANIA, a miniature tethered helicopter having a single, variable pitch rotor [96, 94, 95]. This 0.8-kg rotorcraft had only 3 degrees of freedom (forward and upward motion and pitch). Mounted at the tip of a light whirling arm, the robot lifted itself by increasing its rotor collective pitch. By remotely inclining the servovane located in the propeller wake, the operator caused the helicopter to pitch forward by a few degrees and thus gain speed (up to 6 m/s). FANIA was equipped with an accelerometer and a forward-ventral looking eye with a resolution of only 20 pixels and their corresponding 19 EMDs. FANIA avoided the terrain by increasing its collective pitch as a function of the fused signals transmitted from its eye. Tests in the circular arena showed FANIA jumping over contrasting obstacles [94, 95].

Upon formalizing the optic flow sensed by a flying creature (Fig. 3), we recently came up with an OF based autopilot that we call the optic flow regulator (Fig. 8), which is little demanding in terms of neural implementation and could be just as appropriate for insects as it would be for aircraft [111, 112, 113, 114, 46]. We built a miniature (100-gram) helicopter, called OCTAVE, and equipped it with this autopilot (Fig. 5b). Tested in its circular arena, OCTAVE is able to perform challenging maneuvers such as terrain following at various speeds and smooth take-off and landing [112, 113, 46], as shown in Fig. 9. OCTAVE was also shown to react sensibly to wind perturbations [112], in much the same way as insects do [14, 46]. In this autopilot (Fig. 8) a ventral EMD [115] continuously measures the ventral OF (Fig. 3b) and compares it to an OF set-point ω_{set} . The error signal ε essentially controls the robot's lift L (for details see [46]) and hence its height h via the heave dynamics, so as to maintain the perceived OF at a constant set-point ω_{set} (hence the name: *OF regulator*). This occurs whatever the robot's current groundspeed V_x , whatever disturbances (such as wind) affect that speed, and whatever disturbances (such as a gently sloping terrain) affect the robot's current ground-height h : the *optic flow regulator* will automatically generate a ground height proportional to the groundspeed (see Fig. 9a,b). Two noteworthy results were obtained in these studies [112, 113, 46]:

1. risky manoeuvres such as automatic takeoff, ground avoidance, terrain following, suitable wind reactions and landing are all successfully performed

as a result of one and the same feedback control loop: the *optic flow regulator*.

2. these challenging manoeuvres are all performed *without explicit knowledge of absolute altitude, ground height, groundspeed, airspeed, ascent (or descent) speed and wind speed*.

This bioinspired autopilot therefore differs strikingly from conventional man-designed autopilots, which need a large number of (bulky, heavy and costly) metric sensors (e.g., a radar-altimeter, a baro-altimeter, a Doppler radar, a laser range finder, a GPS receiver, a Pitot tube, a variometer, etc.) to achieve aircraft altitude hold or speed hold. OCTAVE autopilot's objective is actually not to provide for altitude hold or speed hold. Its primary objective is to adapt the groundheight to the groundspeed at all times so that the robot does not crash. It does so at any groundspeeds, raising the robot in proportion to its current groundspeed – whatever internal or external factors affect the latter – without ever *measuring* or *estimating* groundspeed at all [47].

Along the same line, we recently showed that a holonomous and fully actuated air-vehicle can achieve both *lateral obstacle avoidance* and *cruise control* in a corridor, by means of a *dual OF regulator* [119]. Our OF-based autopilot, called LORA III, was shown (in simulation) to automatically adjust both the robot's groundspeed and its clearance from the walls without any needs for *measuring* groundspeed and distance. In other words, the robot achieves this behaviour although it is completely "unaware" of its own groundspeed, unaware of its clearance from the walls, and unaware of the actual corridor width. The LORA III robot navigates at sight on the sole basis of two parameters, which are the set-points of the two intertwined *OF regulators*: a sideways OF set-point and a forward OF set-point. These two parameters were shown to fully constrain the vehicle's behaviour in a straight or tapered corridor, and to impose either wall-following or centering in the corridor [119].

By performing electrophysiological recordings on *fly-ing* flies with soft microelectrodes, we discovered that the muscle attached to the base of the retina [60] is responsible for a *retinal microscanning process*. The whole retinal mosaic (Fig. 2a) is set into translation by a few micrometers repetitively (at about 5 Hz) underneath the facet mosaic, causing the visual axes to rotate by a few degrees [37]. This finding initiated two major biorobotic projects at our laboratory, which aimed to decipher the advantages of this enigmatic process. In

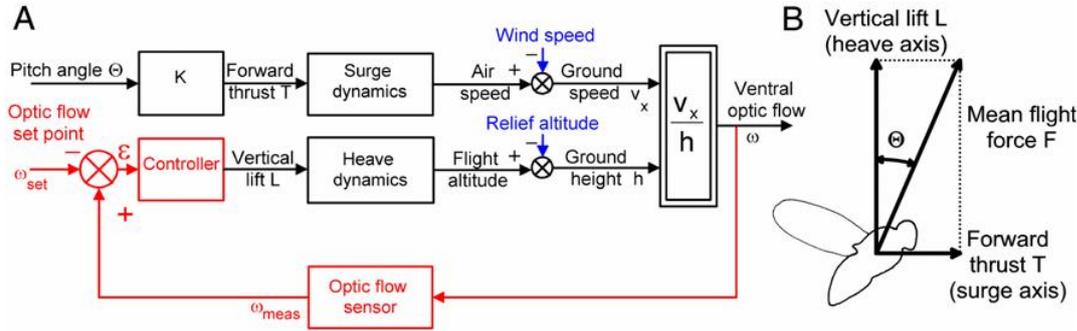


FIGURE 8. Optic Flow Regulator implemented onboard the robot OCTAVE. The control loop adjusts the vertical lift L at all times to maintain the ventral OF, ω constant and equal to the set point (ω_{set}). (a) The upper (open loop) pathway describes how an increase in pitch angle Θ (see b) causes a proportional increase in forward thrust T , and thus in groundspeed V_x . The bottom (red) pathway shows how the ventral OF measured (ω_{meas}) is compared with an OF set-point (ω_{set}). The error signal ($\omega = \omega_{meas} - \omega_{set}$) delivered by the comparator drives a controller adjusting the vertical lift L , and hence the groundheight h , so as to maintain the OF (ω) constant and equal to ω_{set} . The right part of this functional diagram defines the ventral OF: $\omega = V_x/h$, and shows the specific points at which two disturbances impinge on the system: relief altitude and wind speed. (b) Like flies, bees, and helicopters, OCTAVE gains speed by pitching its mean flight force vector forward at a small angle Θ with respect to the vertical (from [46]).

contrast with the robots described above, the two novel robots we built, SCANIA and OSCAR, owe their visual performances to exploiting a purely *rotational* optic flow. Both projects are based on the hypothesis that the microscanning process in flies operates in connection with motion detection.

SCANIA is a 0.7-kg wheeled Cyclopean robot that is able to move about under its own visual control, avoiding the contrasting walls of a square arena despite the low resolution of its eye (only 24 pixels) [90, 89]. This ability is the result of a symmetrical anterograde retinal microscanning process, which assists the robot in *detecting* obstacles located close to the heading direction (i.e., near the frontal “pole” of the optic flow field, defined as $\phi = 0$ in Eq. (1)). The periodic microscanning amounts to periodically adding a given amount of *rotational* optic flow ω_r to the very small *translational* optic flow ω_t generated by frontal obstacles (for which ϕ is small in Eq. (1)). Since the amount of added OF (ω_r) is known onboard, the robot is able to subtract it out subsequently from the overall OF measured, thus recovering the purely *translational* OF (ω_t), i.e., the one OF component that depends on the distance to obstacles (Eq. (1)). Here, the advantage of microscanning is to improve the *detection* of small translational OF that would otherwise have remained subliminal. A similar principle has been taken up again to drive a smaller robot along a contrasting wall and a miniature fly-like retinal microscanner has been built as a MOEMS (micro-opto-

electro-mechanical system) sensor [91, 92, 64].

The second project that we developed on the basis of the fly’s retinal microscanner ended up with a novel optronic sensor, called OSCAR [127], and a novel aerial robot, the robot OSCAR [130, 128, 114] that was equipped with this sensor. The robot OSCAR is attached to a thin, 2-meter long wire secured to the ceiling of the laboratory. It is free to adjust its yaw by driving its two propellers differentially (Fig. 5c). This miniature (100-gram) robot is able to visually fixate a nearby “target” (a dark edge or a bar) thanks to its microscanning process. It is also able to track the target at angular speeds of up to $30^\circ/s$ – a value similar to the maximum tracking speed of the human eye. Target fixation and tracking occur regardless of the distance (up to 2.5 m) and contrast (down to 10%) of the target, and in spite of major disturbances such as pendulum oscillations, ground effects, (gentle) taps and wind gusts. The OSCAR sensor was shown to be able to locate an edge with a resolution of 0.1 degrees, which is 40 times finer than the interreceptor angle $\Delta\phi$ [128]: it has acquired *hyperacuity*, a property that makes the principle appealing for accurate stabilization of various platforms [48], as well as for nonemissive power line detectors on-board fullscale helicopters [131]. Inspired by the fly’s thoracic halteres that have long been known to act as gyroscopes, we equipped the OSCAR robot with an additional *rate control loop* based on a MEMS rate gyro. The interplay of the two (visual and inertial) control loops enhances

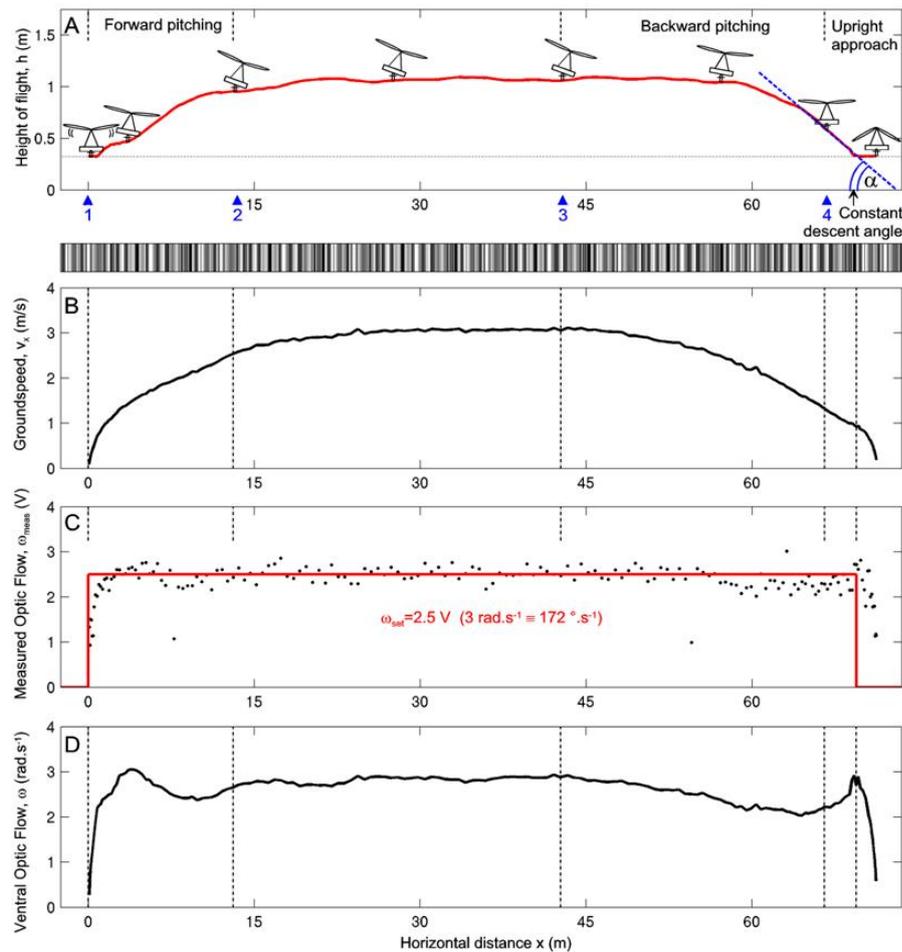


FIGURE 9. Flight parameters monitored during a 70-meter flight (consisting of about 6 laps) performed by the robot OCTAVE (Fig. 5b) equipped with the optic flow regulator shown in Fig. 8. (a) Vertical trajectory in the longitudinal plane. Between arrowheads 1 and 2, the operator pitched OCTAVE forward rampwise by an angle $\Delta\Theta = +10^\circ$. The ensuing increase in groundspeed (see b) caused the robot to climb and fly level at a groundheight of approximately 1 meter. After flying 42 meters, OCTAVE was pitched backwards rampwise by an opposite angle $\Delta\Theta = -10^\circ$ (between arrowheads 3 and 4), and the ensuing deceleration (see b) caused the robot to land smoothly with a constant-descent-angle final approach. (b) Groundspeed V_x monitored throughout the journey. (c) Output ω_{meas} of the OF sensor showing the relatively small deviation from the OF set-point ω_{set} (in red) throughout the journey. (d) Output (true optic flow ω calculated as V_x/h) of the feedback loop. It can be seen that the ventral OF resulting from the MH flight pattern was being held virtually – but not perfectly – constant throughout the journey (from [46]).

both the stability and the dynamic performances of the yaw fixation and pursuit system [130, 128]. More recently we introduced a mechanical *decoupling* between the eye and the body and implemented a *vestibulo-ocular reflex* that was key to maintaining the robot's gaze perfectly fixed on the target [129]. In the latest development, the gaze of the aerial robot OSCAR II remains virtually locked onto the target in spite of drastic thumps

that we deliberately applied to the robot's body by means of a "slapping machine" [70].

5 Biological returns

When a robot mimics specific types of behavior, specific neurons or neural architectures it is likely to provide neuroethologists [136] and neurophysiologists

[137] with useful feedback information. The following are some examples of the biological “returns” that our biorobotic approach has yielded:

1. Monitoring the “optic flow” does not necessarily imply calculating the *distance* to obstacles. While our early terrestrial robots, the Robot-Fly (Fig. 5a) and SCANIA, accessed the distance D to the obstacles according to Eq. (1) (i.e., by knowing θ and V and measuring ω with their EMDs) the flying robots FANIA and OCTAVE go one step further by being able to fly around at relatively high groundspeeds *without any knowledge of the actual distance to the terrain*, in much the same way as honeybees seem to operate [121, 120]. OCTAVE visuomotor control loop acts upon the lift so as to maintain a reference OF at all times, whatever the current groundspeed. The *OF regulator* allows the robot to take-off, avoid ground collisions, follow terrain, react suitably to headwinds and downwinds and land safely, *without any knowledge of local height over terrain, absolute altitude, groundspeed, descent speed and wind speed* [111, 112, 113]. In a recent paper [46] we showed that our OF regulator scheme accounts for many hitherto unexplained findings published during the last 70 years on insects’ visually guided performances; for example, it accounts for the observations that honeybees descend in a headwind [14], land with a constant slope [121, 120], and drown when travelling over mirror-smooth water [61].
2. The small Cyclopean robot SCANIA described in Section 4, which uses a retinal microscanner emulating the one we recently discovered in the fly compound eye, is able to recover the negligibly small *translational* OF associated with objects oriented close to the vehicle’s heading direction. With hindsight, this suggests that the fly may also use its retinal scanner for the same task, eliminating the problem caused by the *frontal pole* of the optic flow field and that caused by the relatively coarse spatial sampling of the eye ($\Delta\phi \approx 1$ to 2°).
3. The aerial robot OSCAR, which is also based on the fly’s microscanner, emits behaviors which are reminiscent of both the hoverfly’s ability to clamp its rotational velocity at zero while hovering near a tree, in wait for passing mates or intruders [23], and the male housefly’s ability to track the female

smoothly during sexual pursuit [141]. We therefore propose that these observed fly behaviors may *result*, at least in part, from the very presence of the microscanner in the compound eye, which is known to come into play only during locomotion, leaving the retina perfectly stationary at rest [37].

Some people tend to feel that computer simulations are all we need [98], and that taking the trouble to *construct* a physical robot is a disproportionately difficult means of checking the soundness of ideas gleaned from nature. There is everything to be gained, however, from completing computer simulations with real physical simulations, for the following reasons:

- Some theories and simulations that look fine on paper (or on the computer screen) may no longer stand under real-world conditions, or may lack robustness to such an extent that they would be useless on-board an animal or a robot.
- Building a robot allows a supposedly understood biological principle to be properly tested in realistically noisy, unstructured and harsh environments.
- At a given level of complexity, it can actually be easier and cheaper to use a real robot in a real environment than to attempt to simulate both the robot and the vagaries of an environment (see also [16, 100]).
- Working with a real robot, with its dynamics, nonlinearities, noisy circuits and other imperfections can help one to decipher the clever tricks evolution has come up with, which can be fathoms apart from what the 21st century scholars were expecting (see also: [101, 93]).

6 Potential applications to manned and unmanned air- and spacecraft

Our insect-inspired robots are minimalistic demonstrators that simply illustrate the fact that biology-derived principles naturally tested for millions of years can be implemented on vehicles at various scales. Further developments along this line may lead to alternatives to the sensors that are in use onboard various terrestrial, aerial or space vehicles.

OF-sensing fits a more general framework called “active perception” [5] where the *ad hoc* movement of a sensor reduces the processing burden involved in perceptual

tasks [6, 1]. Here, it is a *motion sensor* that moves, either as a result of locomotion (in the Robot-Fly, FA-NIA, OCTAVE and LORA III) or as a result of retinal microscanning (in the robots SCANIA and OSCAR).

An OF sensor is a small and lightweight piece of hardware that may be suitable for space applications. It has the advantage to be sensitive to *relative motion*. This could be, for example, the relative motion between a spacecraft and a landing site or between a satellite and its mothership.

An OF sensor is a *passive (nonemissive)* sensor that relies on *photons* and *contrast*. It is therefore no good at night or under poor visibility conditions (fog, dust, rain, snow) and would not compete with *active* sensors such as FLIRs, LADARs or Millimeter-Wave RADARs. But photons and contrast are available in many terrestrial and extraterrestrial worlds such as the Moon, Mars, other planets, asteroids and comets. Should contrast deteriorate in the visible range for whatever reason (e.g., a brownout presented by a dust storm), it may reappear in other portions of the electromagnetic spectrum, between UV and terahertz, so that OF would still be measurable. After all, a suitable shift in spectral sensitivity into the near UV made honeybees able to detect contrast within flowers where no contrast is seen by the human eye. Bees also teach us that it is not advised to work at night in hostile environments, a lesson that may apply to future aerial Martian rovers as well.

Our visually guided robots all make use of self-generated optic flow to carry out humble tasks such as detecting, locating, avoiding or tracking environmental features. Two principles OCTAVE and LORA III may pave the way for automatic piloting without any needs for measuring groundspeed, distance or altitude, thereby eliminating the need for some cumbersome and heavy aerospace sensors. In addition, both OCTAVE and OSCAR autopilots put little constraints on the OF sensor's *range*, owing to the inherent feedback loop of the *OF regulator*, which strives to maintain the OF around a given set-point: only *deviations* from the set-point need to be detected.

In the aerospace context, we anticipate that the bio-inspired OF sensors and visual guidance principles at work in the various robots presented would display the following assets and liabilities:

6.1 Potentials and limitations of OCTAVE and LORA III principles

Both OCTAVE and LORA III could serve to provide better situation awareness for navigation-guidance and control of manned or unmanned vehicles. The OCTAVE autopilot sheds light on the fact that insects may be *permanently unaware of their vertical position* and yet rarely crash. The OF sensor senses the ratio of vertical position to groundspeed and the *OF regulator* makes the altitude proportional to the groundspeed at all times, without the need to *measure* vertical position and groundspeed. Craft's behavior and survival are at a premium, knowledge of speed and distance is of secondary concern. The OCTAVE principle could be harnessed to guide a robotic rover on the surface of the Moon or Mars, or an aerial Martian rover flying low altitude terrain following exploratory missions. Critical, short-lived events like descent and landing of an aerial rover or a space probe could benefit from the OCTAVE principle, since it ensures automatic landing with final approach at constant slope [114, 46]. Flying rovers in Martian canyons and gullies could make use of the LORA III principle, whereby both the craft's groundspeed and its clearance from the walls are constrained by the canyon's width [119]. Each rover could be equipped with its own autonomous guidance system, without relying on any data links with a ground or orbital station. This is an advantage, considering the many potential anomalies of any data links, such as noise and "drop-outs", asynchronization of data transmissions, latency in communication, etc. Similar OF-based principles could probably be applied to safe vehicle spacing in swarms of spacecraft without a centralized control system.

6.2 Potentials and limitations of SCANIA and OSCAR's principles

A lesson from the OSCAR principle is that microscanning is an effective and computer-efficient way to improve image resolution and accuracy in localization. OSCAR provides *hyperacuity* at low cost and low weight, with potential applications to detect-sense-and-avoid systems, fine-guidance sensors and tracking systems, such as novel star trackers that reduce the need to carry other attitude sensors. OSCAR's range of vision will depend on resolution. The OSCAR principle could be applied to automatic craft stabilization with respect to another craft, thus helping autonomous proximity operations such as rendez-vous and docking maneuvers,

e.g., of microsattellites with their mother orbiter. In particular, OSCAR could detect the vanishingly small relative angular velocities between two craft in their final approach. An OSCAR sensor array could be used in synergy with advanced Ladars and video guidance sensors for operations at close range from a spacecraft in refueling or other servicing missions. It could also provide nonemissive detection capability for wires and obstacles during day [131, 48]. During night and low visibility, it would need to fuse its data with those of FLIRs, LADARs or Millimeter-Wave RADARs that are, however, much more power-hungry. As for the OSCAR II principle (involving mechanical decoupling between eye and body [70]), it could serve, e.g., to reject visual perturbations caused by locomotion or wind in exploration rovers.

7 Conclusions

The biorobotic approach that we initiated in 1985 is a transdisciplinary approach which turns out to be most rewarding because it can kill two flies with one stone:

- It can be used to implement a basic principle abstracted from nature and check its soundness, robustness and scaling on a physically embodied machine in a real physical environment. This «reconstruction approach» opens the way to novel devices and machines, particularly in the field of cheap sensory-motor control systems for autonomous vehicles and micro-vehicles, which can thus benefit from the million-century long experience of biological evolution.
- It yields valuable feedback information in the fields of neuroethology and neurophysiology, as it urges us to look at sensory-motor systems from a new angle, raises new biological questions and suggests new experiments to be carried out on insects for iteratively improving our understanding.

Modern space technology, though extremely advanced, might benefit from taking a glance at the way ace pilots such as insects cope with most complex environments. Flying insects are not only the oldest fly-by-wire aircraft, they are able to do much with little: few neurons in the brain, few pixels in the eyes. The diminutive insects no doubt can give us lessons in sensory-motor integration and implementation. We have discussed briefly how the optic flow based principles that we have discovered may be relevant to autonomous

spacecraft behaviors such as obstacle detection, hazards avoidance, tracking, terrain following, landing, docking, etc. Some of these principles could be implemented on-board spacecraft of very small size such as picosatellites, with a view to reduce launch weight, size and power consumption.

Acknowledgments

I am very grateful to numerous colleagues who have worked at the Laboratory over the years, with whom I have had many stimulating discussions: C. Blanes, J. M. Pichon, N. Martin, F. Mura, T. Netter, S. Viollet, F. Ruffier, S. Amic, J. Serres, D. Dray, L. Kerhuel and M. Boyron. This research has been supported by CNRS (Life Sciences, Engineering Sciences, Information Science and Technology, Cognitive Science and Microsystem Programs), and by various EU contracts (ESPRIT, TMR and IST-1999-29043)

References

- [1] Y. Aloimonos. *Active Perception*. Lawrence Erlbaum Associates, Inc. Mahwah, NJ, USA, 1993.
- [2] R. Arkin. *Behavior-Based Robotics*. MIT Press, 1998.
- [3] F. Aubépart and N. Franceschini. Bio-inspired optic flow sensors based on FPGA: Application to Micro-Air-Vehicles. *Microprocessors and Microsystems*, 31(6):408–419, 2007.
- [4] J. Ayers, J. Davis, and A. Rudolph. *Neurotechnology for Biomimetic Robots*. MIT Press, 2002.
- [5] R. Bajcsy. Active perception versus passive perception. In *Proceedings of the 3rd IEEE Workshop Comput. Vis.: Representation and control*, volume 8, pages 55–59, Bellaire, Mich. U.S.A., 1985.
- [6] D. Ballard. Animate vision. *Artificial Intelligence*, 48(1):57–86, 1991.
- [7] H. B. Barlow, J. P. Frisby, A. Horridge, and M. Jeeves. *Natural and artificial low-level seeing systems*. Oxford University Press, New York, 1993.
- [8] F. G. Barth, J. A. C. Humphrey, and T. W. Scomb. *Sensors and Sensing in Biology and Engineering*. Springer, 2003.

- [9] C. Blanes. Appareil visuel élémentaire pour la navigation à vue d'un robot mobile autonome. *MS in Neuroscience, University of Aix-Marseille II, Marseille, France*, 1986.
- [10] C. Blanes. *Guidage visuel d'un robot mobile autonome d'inspiration bionique. 2 ème partie: implémentation opto-électronique et réalisation du prototype*. Thesis work at the Neurocybernetics Res. Group, Marseille, National Polytechnic Institute, Grenoble, 1991.
- [11] A. Borst and J. Haag. Neural networks in the cockpit of the fly. *Journal of Comparative Physiology A: Sensory, Neural, and Behavioral Physiology*, 188(6):419–437, 2002.
- [12] V. Braitenberg. Patterns of projection in the visual system of the fly. I. Retina-lamina projections. *Experimental Brain Research*, 3(3):271–298, 1967.
- [13] V. Braitenberg. *Vehicles*. MIT Press, 1984.
- [14] H. D. Bräuninger. Über den Einfluss meteorologischer Faktoren auf die Entfernungsweisung im Tanz der Bienen. *Journal of Comparative Physiology A: Sensory, Neural, and Behavioral Physiology*, 48(1):1–130, 1964.
- [15] A. D. Briscoe, G. D. Bernard, A. S. Szeto, L. M. Nagy, and R. H. White. Not all butterfly eyes are created equal: Rhodopsin absorption spectra, molecular identification, and localization of ultraviolet-, blue-, and green-sensitive rhodopsin-encoding mRNAs in the retina of *Vanessa cardui*. *The Journal of Comparative Neurology*, 458(4):334–349, 2003.
- [16] R. A. Brooks. *Cambrian intelligence*. MIT Press, 1999.
- [17] E. Buchner. Behavioural analysis of spatial vision in insects. In M. Ali, editor, *Photoreception and Vision in Invertebrates*, pages 561–621. Plenum, New-York, 1984.
- [18] M. Burrows. *The neurobiology of an insect brain*. Oxford University Press, 1996.
- [19] C. Chang and P. Gaudiano. Biomimetic Robotics. *Robotics and Autonomous Systems*, 30(1):1–2, 2000.
- [20] D. Cliff, P. Husbands, J. A. Meyer, and S. W. Wilson. *From Animals to Animats 3: Proceedings of the Third International Conference on Simulation of Adaptive Behavior*. MIT Press, 1994.
- [21] T. Collett. Peering: A locust behaviour pattern for obtaining parallax information. *Journal of Experimental Biology*, 120(11):237–241, 1978.
- [22] T. Collett, H. O. Nalbach, and H. Wagner. Visual stabilization in arthropods. In F. A. Miles and J. Wallman, editors, *Visual motion and its role in the stabilisation of gaze*, volume 5 of *Reviews of Oculomotor Research*, chapter 11, pages 239–63. Elsevier, 1993.
- [23] T. S. Collett and M. F. Land. Visual control of flight behaviour in the hoverfly *Syrirta pipiens* L. *Journal of Comparative Physiology A: Sensory, Neural, and Behavioral Physiology*, 99(1):1–66, 1975.
- [24] C. M. Comer and R. M. Robertson. Identified nerve cells and insect behavior. *Progress in Neurobiology*, 63(4):409–439, 2001.
- [25] D. Coombs and K. Roberts. 'Bee-bot': using peripheral optical flow to avoid obstacles. In *Proceedings Intelligent robots and computer vision XI*, volume 1835, Bellingham, U.S.A., 1992. SPIE.
- [26] M. Dickinson, L. Tammero, and M. Tarsino. Sensory fusion in free-flight search behavior of fruit flies. In J. Ayers, J. Davis, and A. Rudolph, editors, *Neurotechnology for Biomimetic Robots*, pages 573–592. MIT Press, 2002.
- [27] J. K. Douglass and N. J. Strausfeld. Anatomical organization of retinotopic motion-sensitive pathways in the optic lobes of flies. *Microscopy Research and Technique*, 62(2):132–150, 2003.
- [28] M. Egelhaaf and R. Kern. Vision in flying insects. *Current Opinion in Neurobiology*, 12(6):699–706, 2002.
- [29] N. Franceschini. Pupil and pseudopupil in the compound eye of *Drosophila*. In R. Wehner, editor, *Information Processing in the Visual Systems of Arthropods*, pages 75–82. Springer, Berlin, 1972.
- [30] N. Franceschini. Sampling of the visual environment by the compound eye of the fly: fundamentals and applications. In A. Snyder and R. Menzel, editors, *Photoreceptor Optics*, pages 98–125. Springer, Berlin, 1975.

- [31] N. Franceschini. Chromatic organization and sexual dimorphism of the fly retinal mosaic. *Photoreceptors*. Plenum, New York, pages 319–350, 1984.
- [32] N. Franceschini. Early processing of colour and motion in a mosaic visual system. *Neurosci Res Suppl*, 2:S17–49, 1985.
- [33] N. Franceschini. Sequence-discriminating neural network in the eye of the fly. In F. H. K. Eeckman, editor, *Analysis and Modeling of Neural Systems*, pages 142–150. Kluwer Acad. Pub., Norwell, USA, 1992.
- [34] N. Franceschini. Engineering applications of small brains. *Future Electron Devices Journal*, 7:38–52, 1996.
- [35] N. Franceschini. Visual guidance based on optic flow: a biorobotic approach. *Journal of Physiology-Paris*, 98(1-3):281–292, 2004.
- [36] N. Franceschini, C. Blanes, and L. Oufar. Passive non-contact optical velocity sensor. In *Dossier technique ANVAR/DVAR*, volume 51549. 1986.
- [37] N. Franceschini and R. Chagneux. Repetitive scanning in the fly compound eye. In *Goettingen Neurobiology Report 1997*, page 279. Georg Thieme Verlag, 1997.
- [38] N. Franceschini, R. Hardie, W. Ribi, and K. Kirschfeld. Sexual dimorphism in a photoreceptor. *Nature*, 291(5812):241–244, 1981.
- [39] N. Franceschini and K. Kirschfeld. Etude optique in vivo des éléments photorécepteurs dans l'œil composé de *Drosophila*. *Biological Cybernetics*, 8(1):1–13, 1971.
- [40] N. Franceschini and K. Kirschfeld. Les phénomènes de pseudopupille dans l'œil composé de *Drosophila*. *Kybernetik*, 9:159–182, 1971.
- [41] N. Franceschini, K. Kirschfeld, and B. Minke. Fluorescence of photoreceptor cells observed in vivo. *Science*, 213(4513):1264–1267, 1981.
- [42] N. Franceschini, J. Pichon, and C. Blanes. Bionics of Visuo-Motor Control. In T. Gomi, editor, *Evolutionary Robotics: from intelligent robots to artificial life*, pages 49–67. AAAI Books, Ontario, Canada, 1997.
- [43] N. Franceschini, J. M. Pichon, and C. Blanes. Real time visuomotor control: from flies to robots. In *IEEE Conf. on Advanced Robotics (ICAR 91)*, pages 931–935, Pisa, Italy, 1991.
- [44] N. Franceschini, J. M. Pichon, and C. Blanes. From insect vision to robot vision. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 337(1281):283–294, 1992.
- [45] N. Franceschini, A. Riehle, and A. Le Nestour. Directionally selective motion detection by insect neurons. In D. G. Stavenga and R. C. Hardie, editors, *Facets of Vision*, pages 360–390. Springer, Berlin, 1989.
- [46] N. Franceschini, F. Ruffier, and J. Serres. A Bio-Inspired Flying Robot Sheds Light on Insect Piloting Abilities. *Current Biology*, 17(4):329–335, 2007.
- [47] N. Franceschini, F. Ruffier, S. Viollet, and M. Boyron. Steering aid system for altitude and horizontal speed, perpendicular to the vertical, of an aircraft and aircraft equipped therewith, Dec. 20 2005. Patent PCT/FR2005/11536.
- [48] N. Franceschini, S. Viollet, and M. Boyron. Method and device for hyperacute detection of an essentially rectilinear contrast edge and system for fine following and fixing of said contrast edge, Jan. 3 2003. PCT Patent FR2003/ 002611.
- [49] J. J. Gibson. *The Perception of the Visual World*. Houghton Mifflin, 1950.
- [50] K. G. Götz. Flight control in *Drosophila* by visual perception of motion. *Biological Cybernetics*, 4(6):199–208, 1968.
- [51] M. Goulet, R. Campan, and M. Lambin. The visual perception of relative distances in the wood-cricket, *Nemobius sylvestris*. *Physiological Entomology*, 6(4):357–367, 1981.
- [52] A. Guillot and J. A. Meyer. Biomimetic robotics. *Robotics and Autonomous Systems*, 50(4):129–230, 2005.

- [53] R. Hardie. Functional organization of the fly retina. In D. Ottosson, editor, *Progress in Sensory Physiology*, volume 5, pages 2–79. Springer, Berlin, 1985.
- [54] R. Hardie, N. Franceschini, and P. McIntyre. Electrophysiological analysis of fly retina II: Spectral and polarisation sensitivity in R7 and R8. *Journal of Comparative Physiology A: Sensory, Neural, and Behavioral Physiology*, 133(1):23–39, 1979.
- [55] R. R. Harrison and C. Koch. A Silicon Implementation of the Fly’s Optomotor Control System. In *Neural Computation*, volume 12, pages 2291–2304. MIT Press, 2000.
- [56] K. Hausen. Monocular and binocular computation of motion in the lobula plate of the fly. *Verh Dtsch Zool Ges*, 74(4):70, 1981.
- [57] K. Hausen. Motion sensitive interneurons in the optomotor system of the fly II/The horizontal cells: receptive field organization and response characteristics. *Biological Cybernetics*, 45(2):143–156, 1982.
- [58] K. Hausen and M. Egelhaaf. Neural mechanisms of visual course control in insects. In R. C. Hardie and D. G. Stavenga, editors, *Facets of Vision*, pages 391–424. Springer, Berlin, 1989.
- [59] M. Heisenberg and R. Wolf. *Vision in Drosophila*. Springer-Verlag New York, 1984.
- [60] R. Hengstenberg. Das Augenmuskelsystem der Stubenfliege *Musca domestica*. *Biological Cybernetics*, 9(2):56–77, 1971.
- [61] H. Heran and M. Lindauer. Windkompensation und Seitenwindkorrektur der Bienen beim Flug über Wasser. *Journal of Comparative Physiology A: Sensory, Neural, and Behavioral Physiology*, 47(1):39–55, 1963.
- [62] C. M. Higgins, J. K. Douglass, and N. J. Strausfeld. The computational basis of an identified neuronal circuit for elementary motion detection in dipterous insects. *Visual Neuroscience*, 21(04):567–586, 2004.
- [63] G. Horridge. The Evolution of Visual Processing and the Construction of Seeing Systems. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, 230(1260):279–292, 1987.
- [64] K. Hoshino, F. Mura, and I. Shimoyama. A one-chip scanning retina with an integrated micromechanical scanning actuator. *Journal of Microelectromechanical Systems*, 10(4), 2001.
- [65] G. Hoyle. *Identified neurons and behavior of arthropods*. Plenum Press, New York, 1977.
- [66] S. A. Huber, M. O. Franz, and H. H. Bülthoff. On robots and flies: Modeling the visual orientation behavior of flies. *Robotics and Autonomous Systems*, 29(4):227–242, 1999.
- [67] M. Ichikawa, H. Yamada, and J. Takeuchi. Flying Robot with Biologically Inspired Vision. *J. Robotics and Mechatronics*, 13(6):621–624, 2001.
- [68] F. Iida and D. Lambrinos. Navigation in an autonomous flying robot by using a biologically inspired visual odometer. In G. T. McKee and P. S. Schenker, editors, *Sensor fusion and decentralized control in robotic systems III*, volume 4196, page 86, Bellingham, USA, 2000. SPIE.
- [69] G. Indiveri, J. Kramer, and C. Koch. System implementations of analog VLSI velocity sensors. *IEEE Micro*, 16(5):40–49, 1996.
- [70] L. Kerhuel, S. Viollet, and N. Franceschini. A sighted aerial robot with fast gaze and heading stabilization. In *IEEE Intern. Conf. on Intelligent Robots and Systems (IROS 07)*, pages 2634–2641, San Diego, USA, 2007.
- [71] W. H. Kirchner and M. V. Srinivasan. Freely flying honeybees use image motion to estimate object distance. *Naturwissenschaften*, 76(6):281–282, 1989.
- [72] K. Kirschfeld. Die Projektion der optischen Umwelt auf das Raster der Rhabdomere im Komplexauge von *Musca*. *Exp. Brain Res*, 3:248–270, 1967.
- [73] K. Kirschfeld. The visual system of *Musca*: studies on optics, structure and function. In R. Wehner, editor, *Information Processing in the Visual Systems of Arthropods*, pages 61–74. Springer, Berlin, 1972.
- [74] K. Kirschfeld, R. Feiler, and N. Franceschini. A photostable pigment within the rhabdomere

- of fly photoreceptors N. 7. *Journal of Comparative Physiology A: Sensory, Neural, and Behavioral Physiology*, 125(3):275–284, 1978.
- [75] K. Kirschfeld and N. Franceschini. Optische Eigenschaften der Ommatidien im Komplexauge von Musca. *Biological Cybernetics*, 5(2):47–52, 1968.
- [76] K. Kirschfeld, N. Franceschini, and B. Minke. Evidence for a sensitising pigment in fly photoreceptors. *Nature*, 269(5627):386–390, 1977.
- [77] J. J. Koenderink. Optic flow. *Vision Res*, 26(1):161–79, 1986.
- [78] H. G. Krapp, B. Hengstenberg, and R. Hengstenberg. Dendritic Structure and Receptive-Field Organization of Optic Flow Processing Interneurons in the Fly. *Journal of Neurophysiology*, 79(4):1902–1917, 1998.
- [79] M. Kuwabara and K. Naka. Response of a single retinula cell to polarized light. *Nature*, 184(4684):455–456, 1959.
- [80] D. Lee and H. Kalmus. The Optic Flow Field: The Foundation of Vision. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences (1934–1990)*, 290(1038):169–179, 1980.
- [81] M. Lehrer, M. V. Srinivasan, S. W. Zhang, and G. A. Horridge. Motion cues provide the bee's visual world with a third dimension. *Nature*, 332(6162):356–357, 1988.
- [82] M. A. Lewis and M. A. Arbib. Introduction to the Special Issue on Biomorphic Robots. *Autonomous Robots*, 7(3):207–209, 1999.
- [83] M. A. Lewis and M. E. Nelson. Look before you leap: Peering behavior for depth perception. In R. Pfeifer, B. Blumberg, J. Meyer, and S. Wilson, S, editors, *From Animals to Animats 5: Proceedings of the Fifth International Conference on Simulation of Adaptive Behavior*. MIT Press, 1998.
- [84] P. Maes. *Designing Autonomous Agents: Theory and Practice from Biology to Engineering and Back*. MIT Press, 1990.
- [85] N. Martin and N. Franceschini. Obstacle avoidance and speed control in a mobile vehicle equipped with a compound eye. In I. Masaki, editor, *Intelligent Vehicles*, pages 381–386. MIT Press, 1994.
- [86] M. J. Mataric. Navigating with a rat brain: a neurobiologically-inspired model for robot spatial representation. In J. A. Meyer and S. W. Wilson, editors, *From animals to animats*, pages 169–175. MIT Press, 1990.
- [87] R. Möller, D. Lambrinos, R. Pfeifer, T. Labhart, and R. Wehner. Modeling ant navigation with an autonomous agent. In R. Pfeifer, B. Blumberg, J. Meyer, and S. W. Wilson, editors, *From Animals to Animats 5*, pages 185–194. MIT Press, 1998.
- [88] F. Mura and N. Franceschini. Visual control of altitude and speed in a flying agent. In *From Animals to Animats*, pages 91–99. MIT Press, 1994.
- [89] F. Mura and N. Franceschini. Biologically inspired 'retinal scanning' enhances motion perception of a mobile robot. In A. Bourjault and S. Hata, editors, *1st Europe-Asia Congress on Mechatronics*, volume 3, pages 934–940. ENSM, 1996.
- [90] F. Mura and N. Franceschini. Obstacle avoidance in a terrestrial mobile robot provided with a scanning retina. In M. Aoki and I. Masaki, editors, *Intelligent Vehicles II*, pages 47–52. 1996.
- [91] F. Mura, N. Franceschini, I. Shimoyama, and H. Miura. Retinal scanning improves the estimation of the optic flow for visually-guided navigation. In *14th Conf. Robotics Soc. of Japan.*, pages 1191–1192, Niigata, 1996.
- [92] F. Mura and I. Shimoyama. Visual guidance of a small mobile robot using active, biologically-inspired, eye movements. In *IEEE Intern. Conf. on Robotics and Automation (ICRA 98)*, volume 3, pages 1859–1864, Leuven, Belgium, 1998.
- [93] W. Nachtigall. *Bionik: Grundlagen und Beispiele für Ingenieure und Naturwissenschaftler*. Springer, 2002.
- [94] T. Netter and N. Franceschini. A robotic aircraft that follows terrain using a neuromorphic eye.

- In *IEEE/RSJ Intern. Conf. on Intelligent Robots and System (IROS 02)*, volume 1, pages 129–134, EPFL, Lausanne, 2002.
- [95] T. Netter. and N. Franceschini. Neuromorphic motion detection for robotic flight guidance. *The neuromorphic engineer*, 1:8–9, 2004.
- [96] T. Netter and N. H. Franceschini. Neuromorphic optical flow sensing for Nap-of-the-Earth flight. In *Mobile robots XIV*, volume 3838, page 208, Bellingham, USA, 1999. SPIE.
- [97] T. R. Neumann and H. H. Bülthoff. Insect inspired visual control of translatory flight. In *Proceedings of the 6th European Conference on Artificial Life (ECAL 01)*, pages 627–636, Berlin, Germany, 2001. Springer, Berlin.
- [98] T. R. Neumann, S. Huber, and H. H. Bülthoff. Artificial systems as models in biological cybernetics. *Behavioral and Brain Sciences*, 24(06):1071–1072, 2002.
- [99] M. O’Shea and C. Rowell. Complex neural integration and identified interneurons in the locust brain. In G. Hoyle, editor, *Identified neurons and behavior of arthropods*, pages 307–328. Plenum Press, New York, 1977.
- [100] R. Pfeifer, M. Lungarella, and F. Iida. Self-Organization, Embodiment, and Biologically Inspired Robotics. *Science*, 318(5853):1088, 2007.
- [101] R. Pfeifer and C. Scheier. *Understanding intelligence*. MIT Press, 2001.
- [102] F. Pichaud, A. Briscoe, and C. Desplan. Evolution of color vision. *Current Opinion in Neurobiology*, 9(5):622–627, 1999.
- [103] J. M. Pichon, C. Blanes, and N. H. Franceschini. Visual guidance of a mobile robot equipped with a network of self-motion sensors. In W. Wolfe and W. Chun, editors, *Mobile Robots IV*, volume 1195, pages 44–53, Bellingham, USA, 1990. SPIE.
- [104] M. Pudas, S. Viollet, F. Ruffier, A. Kruusing, S. Amic, S. Leppävuori, and N. Franceschini. A miniature bio-inspired optic flow sensor based on low temperature co-fired ceramics (LTCC) technology. *Sensors & Actuators: A. Physical*, 133(1):88–95, 2007.
- [105] X. Qiu, K. A. J. Vanhoutte, D. G. Stavenga, and K. Arikawa. Ommatidial heterogeneity in the compound eye of the male small white butterfly, *Pieris rapae crucivora*. *Cell and Tissue Research*, 307(3):371–379, 2002.
- [106] W. Reichardt. Autokorrelationsauswertung als Funktionsprinzip des Zentralnervensystems. *Z. Naturforsch*, 12:448–457, 1957.
- [107] W. Reichardt. Movement Perception in Insects. In *Processing of optical data by organisms and by machines*, pages 465–493. North-Holland, 1969.
- [108] W. Reichardt. Evaluation of optical motion information by movement detectors. *Journal of Comparative Physiology A*, 161(4):533–547, 1987.
- [109] A. Riehle and N. Franceschini. Motion detection in flies: parametric control over ON-OFF pathways. *Exp. Brain Res.*, 54(2):390–394, 1984.
- [110] F. Rind. Bioinspired sensors: from insect eyes to robot vision. In T. A. Christensen, editor, *Methods in insect sensory neuroscience*, pages 213–234. 2004.
- [111] F. Ruffier and N. Franceschini. OCTAVE: a bioinspired visuo-motor control system for the guidance of micro-air-vehicles. In A. Gabriel-Vasquez, D. Abbott, and R. Carmona, editors, *Bioengineered and bioinspired systems*, volume 5119, pages 1–12. SPIE, 2003.
- [112] F. Ruffier and N. Franceschini. Visually guided micro-aerial vehicle: automatic take off, terrain following, landing and wind reaction. In *Proceedings IEEE Intern. Conf. on Robotics and Automation (ICRA04)*, pages 2339–2346, New Orleans, USA, 2004.
- [113] F. Ruffier and N. Franceschini. Optic flow regulation: the key to aircraft automatic guidance. *Robotics and Autonomous Systems*, 50(4):177–194, 2005.
- [114] F. Ruffier, J. Serres, and N. Franceschini. Automatic landing and take-off at constant slope

- without terrestrial aids. In *31th European Rotorcraft Forum (ERF05)*, volume 71, pages 1–9, New Orleans, USA, 2005.
- [115] F. Ruffier, S. Viollet, S. Amic, and N. Franceschini. Bio-inspired optical flow circuits for the visual guidance of micro-air vehicles. In *Proceeding of IEEE Intern. Symposium on Circuits And Systems (ISCAS 03)*, volume 3, pages 846–849, Bangkok, Thailand, 2003.
- [116] G. Sandini, J. Santos-Victor, F. Curotto, and S. Garibaldi. Robotic bees. In *Proceedings of the 1993 IEEE/RSJ Intern. Conf. on Intelligent Robots and Systems (IROS93)*, volume 1, New York, USA, 1993.
- [117] R. Sarpeshkar, J. Kramer, and C. Koch. Pulse domain neuromorphic integrated circuit for computing motion, July 14 1998. US Patent 5,781,648.
- [118] C. Schilstra and J. H. Van Hateren. Blowfly flight and optic flow. 1. Thorax kinematics and flight dynamics. *Journal of experimental biology*, 202:1481–1490, 1999.
- [119] J. Serres, D. Dray, F. Ruffier, and N. Franceschini. A vision-based autopilot for a miniature air vehicle: joint speed control and lateral obstacle avoidance. *Autonomous Robots*, 25(1):103–122, 2008.
- [120] M. Srinivasan, S. Zhang, J. Chahl, E. Barth, and S. Venkatesh. How honeybees make grazing landings on flat surfaces. *Biological Cybernetics*, 83(3):171–183, 2000.
- [121] M. V. Srinivasan, S. W. Zhang, M. Lehrer, and T. S. Collett. Honeybee navigation en route to the goal: visual flight control and odometry. *Journal of Experimental Biology*, 199(1):237–244, 1996.
- [122] V. V. Srinivasan and S. Venkatesh. *From Living Eyes to Seeing Machines*. Oxford University Press, 1997.
- [123] N. Strausfeld. Beneath the compound eye: neuroanatomical analysis and physiological correlates in the study of insect vision. In D. G. Stavenga and R. C. Hardie, editors, *Facets of vision*. Springer, Berlin Heidelberg New York, chapter 16, pages 317–359. Springer, Berlin, 1989.
- [124] N. J. Strausfeld. *Atlas of an insect brain*. Springer-Verlag, Berlin, 1976.
- [125] L. F. Tammero and M. H. Dickinson. The influence of visual landscape on the free flight behavior of the fruit fly *Drosophila melanogaster*. *Journal of Experimental Biology*, 205(3):327–343, 2002.
- [126] S. Ullman. Analysis of visual motion by biological and computer systems. *Computer*, 14(8):57–69, 1981.
- [127] S. Viollet and N. Franceschini. Biologically-inspired visual scanning sensor for stabilization and tracking. In *IEEE Intern. Conf. on Intelligent Robots and Systems (IROS 99)*, volume 1, pages 204–209, Kyon-gyu, Korea, 1999.
- [128] S. Viollet and N. Franceschini. Super-accurate visual control of an aerial minirobot. In U. Rückert, J. Sitte, and U. Witkowski, editors, *Autonomous Minirobots for research and edutainment (AMIRE 01)*, pages 215–224, Paderborn, Germany, 2001.
- [129] S. Viollet and N. Franceschini. A high speed gaze control system based on the Vestibulo-Ocular Reflex. *Robotics and Autonomous Systems*, 50(4):147–161, 2005.
- [130] S. Viollet and N. H. Franceschini. Visual servo system based on a biologically inspired scanning sensor. In *Proceedings of SPIE: Sensor fusion and decentralized control*, volume 3839, page 144, Bellingham, USA, 1999. SPIE.
- [131] S. Viollet, J. Michelis, and N. Franceschini. Toward a bio-inspired nonemissive powerline detection system. In *Proc. 32th European Rotorcraft Forum*, number AD09, Maastricht, 2006. ERF.
- [132] H. Wagner. Flow-field variables trigger landing in flies. *Nature*, 297(5862):147–148, 1982.
- [133] H. Wagner. Flight Performance and Visual Control of Flight of the Free-Flying Housefly (*Musca Domestica* L.) I. Organization of the Flight Motor. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 312(1158):527–551, 1986.

- [134] M. Wakakuwa, M. Kurasawa, M. Giurfa, and K. Arikawa. Spectral heterogeneity of honeybee ommatidia. *Naturwissenschaften*, 92(10):464–467, 2005.
- [135] J. Wallman and F. A. Miles. *Visual Motion and its Role in the Stabilization of Gaze*. Elsevier Amsterdam, 1993.
- [136] B. Webb. Can robots make good models of biological behaviour? *Behavioral and Brain Sciences*, 24(06):1033–1050, 2002.
- [137] B. Webb. Robots in invertebrate neuroscience. *Nature*, 417:359–363, 2002.
- [138] B. Webb and T. Consi. *Biorobotics*. MIT Press, 2001.
- [139] B. Webb, R. R. Harrison, and M. A. Willis. Sensorimotor control of navigation in arthropod and artificial systems. *Arthropod Structure and Development*, 33(3):301–329, 2004.
- [140] R. Wehner. *Spatial vision in arthropods*, volume VII/6C, pages 288–616. Springer, Berlin, 1981.
- [141] C. Wehrhahn. Sex-specific differences in the chasing behaviour of houseflies (*Musca*). *Biological Cybernetics*, 32(4):239–241, 1979.
- [142] T. Whiteside and G. Samuel. Blur zone. *Nature*, 225(5227):94–95, 1970.
- [143] M. Zaretsky and C. H. F. Rowell. Saccadic suppression by corollary discharge in the locust. *Nature*, 280(5723):583–585, 1979.
- [144] J. C. Zufferey and D. Floreano. Toward 30-gram Autonomous Indoor Aircraft: Vision-based Obstacle Avoidance and Altitude Control. In *IEEE Intern. Conf. on Robotics and Automation (ICRA 05)*, pages 2594–2599, Barcelona, Spain, 2005.



Acta Futura 3 (2009) 35-42

DOI: 10.2420/AF03.2008.29

**Acta
Futura**

The Concept of PEIS-Ecology: Integrating Robots in Smart Environments

ALESSANDRO SAFFIOTTI*

AASS Mobile Robotics Lab. Department of Technology, University of Örebro, Sweden

Abstract. The concept of Ecology of Physically Embedded Intelligent Systems, or PEIS-Ecology, combines insights from the fields of ubiquitous robotics and ambient intelligence to provide a new solution to building intelligent robots in the service of people. In this note I introduce this concept, summarize its main technological aspects, and speculate on its potential impact for space exploration.

1 Introduction

Autonomous robotic devices are vital to the performance of space missions. These range from smart servo-controlled devices (e.g., an antenna or a camera) to autonomous mobile robots (e.g., planetary rovers) and to future interactive astronaut assistants. Each robotic device operates in its specific environment, which is usually not fully predictable and not fully observable, and which may include humans. In the classical view of autonomous robotics, the robot and its environment are seen as two distinct entities. This view is often assimilated to a two-player antagonistic game, in which the robot has to find a strategy to achieve its goals in spite of the actions performed by the environment.

The “PEIS-Ecology” approach to autonomous robotics takes an ecological view of the robot-

environment relationship, in which the robot and the environment are seen as parts of the same system, engaged in a symbiotic relationship toward the achievement of a common goal, or equilibrium status. We assume that robotic devices (or PEIS, for “Physically Embedded Intelligent Systems”) are pervasively distributed throughout the working space in the form of sensors, actuators, smart appliances, RFID-tagged objects, or more traditional mobile robots; and that these PEIS can communicate and collaborate with each other by providing information and by performing actions. Humans can also be included in this approach as another species of PEIS inside the same ecosystem.

The PEIS-Ecology approach was developed in the context of a collaborative research project between Sweden and Korea between 2004 and 2008. The target application of this approach was everyday domestic assistance, especially (but not only) to elderly people. However, this approach can in principle be applied to other domains, both indoor and outdoor, including space exploration. In this note, we outline the concept of PEIS-Ecology, discuss the major scientific and technological challenges entailed by its realization, and show the results achieved in the course of our development with respect to these challenges. We also discuss the expected developments of the PEIS-Ecology approach in the near future, and speculate on the applicability and benefits of this approach in the context of space exploration.

*E-mail address: asaffio@aass.oru.se

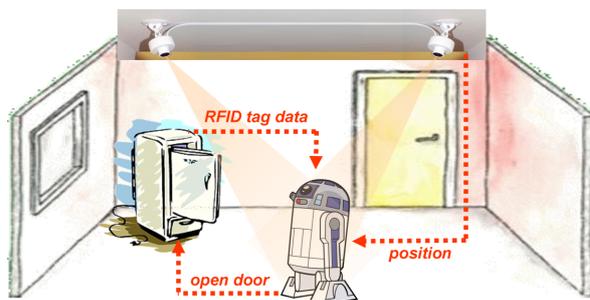


FIGURE 1. A simple example of PEIS-Ecology.

2 Current Status

The PEIS-Ecology project was started in October 2004 as a collaborative effort between Sweden (the AASS Mobile Robotics Lab at the University of Örebro) and Korea (the Electronic and Telecommunication Research Institute). Currently this project has attracted additional funding from the Swedish Research Council and from the Swedish National Graduate School in Computer Science, and it has grown to an effort of about 80 person-month per year. In this section, we briefly recall the main concept underlying this project, and summarize the main scientific and technological achievements in the first 40 months of this project.

2.1 The Concept of PEIS-Ecology

The concept of PEIS-Ecology, originally introduced by Saffiotti and Broxvall [15], builds upon the following ingredients:

First, any robot in the environment is abstracted by the *uniform notion* of PEIS¹ (Physically Embedded Intelligent System). The term “robot” is taken here in its most general interpretation: any device incorporating some computational and communication resources, and able to interact with the environment via sensors and/or actuators. A PEIS can be as simple as a toaster or as complex as a humanoid robot. In general, we define a PEIS to be a set of inter-connected software *components* residing in one physical entity. Each component can be connected to sensors and actuators in that physical entity, as well as to other components in the same PEIS or in other PEIS.

Second, all PEIS are connected by a *uniform communication model*, which allows the exchange of informa-

tion among the individual PEIS-components, while hiding the heterogeneity of the PEIS and of the physical communication layers. In practice, we use a distributed communication model that combines a tuple-space with an event mechanism.

Third, all PEIS in an ecology can cooperate by a *uniform cooperation model*, based on the notion of linking functional components: each participating PEIS can use functionalities from other PEIS in the ecology to complement its own. Functionalities here are meant to be modules that produce and consume information, and may interact with the physical environment by means of sensors and actuators. Typically, functionalities are in one-to-one correspondence to the software components in a PEIS.

Finally, we define a PEIS-Ecology to be a collection of inter-connected PEIS, all embedded in the same physical environment.

As an illustration of these concepts, consider a home robot with the task of grasping a milk bottle from the fridge. (See Figure 1.) In a classical approach, the robot would use its sensors to acquire information from the environment — e.g., to self-localize, and to acquire the relevant parameters of the fridge handle and of the milk bottle. It would then use its actuators to manipulate the environment — e.g., to open the fridge door and to grasp the milk bottle. In a PEIS-Ecology, by contrast, the robot would ask (some of) the needed information from the environment — e.g., it would get its position from cameras in the ceiling; and it would get the shape, weight, and grasping points of the milk bottle from the bottle itself, equipped with a mote or an RFID tag. It would also ask the environment to perform (some of) the needed actions — e.g., it would ask the fridge to open its door.

Given a PEIS-Ecology, we call a set of connections between components within and across the ecology a *configuration* of that PEIS-Ecology. Importantly, the same ecology can be configured in many different ways depending on the context — e.g., depending on the current task, the environmental situation, and available resources. In the above example, if the robot exits the field of view of the ceiling cameras, then the ecology may be reconfigured to let it use another PEIS in the ecology for localization, e.g., a camera on the fridge or the robot’s own odometric system.

A PEIS-Ecology redefines the very notion of a *robot* to encompass the entire environment: a PEIS-Ecology may be seen as a “cognitive robotic environment” in which perception, actuation, memory, and processing are per-

¹PEIS is pronounced /peIs/ like in “pace”.

vasively distributed in the environment. The complex functionalities of this environment are not determined in a centralized way, but they emerge from the cooperation of many simpler, specialized, ubiquitous PEIS devices. The number and capabilities of these devices do not need to be known *a priori*: new PEIS can join or leave the ecology at any moment, and their existence and capabilities should be automatically detected by the other PEIS.

The PEIS-Ecology approach simplifies many of the difficult problems of current autonomous robotics by replacing complex on-board functionalities with simple off-board functionalities plus communication. In the milk example above, the global localization of the robot is easily achieved by the static cameras; and the best way to access the properties of the milk box is to store those properties in the box itself. The PEIS-Ecology approach can also help us to address problems which are beyond the capabilities of current robotic systems. An example of this type is reported in [1], where a PEIS-Ecology is used to solve a home monitoring task involving the use of olfaction, which would be hard to solve otherwise due the current limitations of mobile olfaction.

The PEIS-Ecology approach can also bring a number of pragmatic benefits. A PEIS-Ecology is intrinsically modular, flexible and customizable. Users would only need to acquire new robotic components as needed, e.g., starting with just a simple robotic vacuum cleaner and adding new PEIS devices according to their changing needs and desires. Thus, the PEIS-Ecology approach is likely to provide an affordable and acceptable road to include robotic technologies in everyday environments. Since each new PEIS can combine its functionalities with those of the already existing ones, the value of the whole PEIS-Ecology can increase more than linearly with its cost.

In spite of the above attractive benefits, the development of PEIS-Ecology entails a number of new research challenges that need to be solved before this potential can be fully exploited. The next three subsections outline three of the most peculiar ones, together with our progress in addressing them.

2.2 Dealing with Heterogeneity

The first fundamental challenge is heterogeneity. A PEIS-Ecology may include highly heterogeneous devices, which rely on different hardware and software platforms and different communication media. Heterogeneity may also arise from the different levels at

which the devices need to exchange information: from raw data streams to oneshot data readings to symbolic communication. In face of this, a PEIS-Ecology should provide the means to establish a meaningful communication between different PEIS. Physical diversity should be abstracted, and contents should make reference to a common ontology and measurement system. Achieving this requires a suitable middleware.

In our work, we have developed an open-source middleware called the PEIS-kernel [2]. This provides uniform communication primitives, and performs services like network discovery and routing of messages between PEIS on a P2P network. The PEIS-kernel also implements a communication model based on a distributed tuple-space, endowed with the usual `insert` and `read` operations. In addition, it provides event-based primitives `subscribe` and `unsubscribe`, by which a PEIS-component can signal its interest in a given tuple key. When an insert operation is performed, all subscribers are notified. Subscription, notification, and distribution of tuples are managed by the PEIS-kernel in a way which is transparent to the PEIS-component.

The PEIS-kernel can cope with the fact that PEIS may dynamically join and leave the ecology. At any moment, each PEIS-component can detect the presence of other components and trade with them the use of functionalities. For instance, if the navigation component in the robot in Figure 1 above requires a localization functionality, it simply looks for a tuple announcing a *compatible* functionality in any PEIS-component: if one is found, that component is booked and a subscription to it is created. Compatibility is decided using a shared PEIS-Ontology.

The PEIS-Ecology middleware has been released as open-source under a set of GNU licenses, and it is available from the project website [14].

2.3 Integrating the physical and the digital world

The second fundamental challenge of a PEIS-Ecology can be described as follows. In a classical robotic system, the robot's interaction with the environment and its objects is physically mediated: properties of the objects are estimated using sensors, and their state can be modified using actuators. In a PEIS-Ecology, a robot (PEIS) can interact with an object (another PEIS) both physically and digitally: the robot can directly query properties from the object, and it can ask it to perform an action. The new challenge here is how to coordinate and

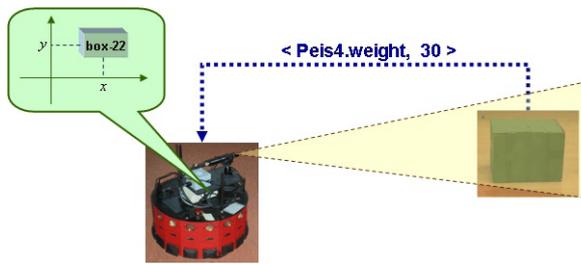


FIGURE 2. Linking perceptual information and digital information in a PEIS-Ecology.

integrate these two forms of interaction.

Consider a robot in a PEIS-Ecology, which is facing a closed door. The robot would need to know if this door can be opened. Suppose that the robot is aware, through the advertisement mechanism mentioned above, that there is a PEIS in the PEIS-Ecology, with ID = `Peis301`, which offers the action 'open'. If the robot can establish that the door which it is perceiving in front of it is the same physical object as `Peis301`, then the robot will know that that door can perform the action of opening, by linking the (digital) action 'open' of `Peis301` with the (physical) action of opening the specific door. It will also know that in order to open that specific door it needs to send the request `<open>` to the PEIS with ID = `Peis301`. A similar linkage would allow the robot to augment its perceptual knowledge about an object by using the properties communicated by the corresponding PEIS.

Our approach to cope with this challenge is based on an extension of the concept of *perceptual anchoring* [4]. Anchoring is the process of connecting, inside an intelligent system, the symbols used to denote an object (e.g., `box-22`) and the percepts originating from the same objects (e.g., a green blob in the camera image). More specifically, the robot queries the tuple-space for all `PhysicalRepresentation` tuples of each PEIS in the ecology (each PEIS must publish this tuple by convention). It then tries to match these tuples to the perceived properties of the object in front of it, e.g., being box-shaped, green, and of a certain size. If the matching succeeds for a given PEIS, say the one with ID = `Peis4`, then the robot can ask additional properties to `Peis4` (e.g., its weight) and combine these properties with the observed ones, e.g., to decide if the box can be pushed.

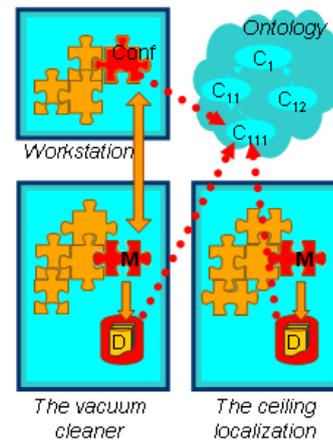


FIGURE 3. Outline of our self-configuration framework.

2.4 Self-configuration

Perhaps the strongest added value of a PEIS-Ecology comes from the ability to integrate the functionalities available in the different PEIS according to a given configuration, and to automatically create and modify this configuration depending on the current context. Here, the relevant contextual conditions include the current task(s), the state of the environment, and the resources available in the ecology. Self-configuration is the key to flexibility, adaptability and robustness of the system — in one word, to its *autonomy*.

The problem of self-configuration is a hard open problem for autonomous systems in general, and for distributed robotic systems in particular. In a PEIS-Ecology, this problem is exacerbated by the fact that a PEIS-Ecology is highly heterogeneous and intrinsically dynamic.

Our current approach to self-configuration is partly inspired by work in the field of web service composition [7]. It is based on the following ingredients.

- An *advertising mechanism* that allows any PEIS to dynamically join the ecology and let all the other PEIS know about the functionalities it can provide.
- A *discovery mechanism* that allows each PEIS to find which other PEIS can provide a functionality compatible with its needs.
- A *configuration mechanism* able to create a configuration for a given task by composing functionalities from different PEIS.

- A *monitoring mechanism* able to change the configuration if these functionalities become unavailable.

The above mechanisms help to cope with the dynamic aspect. To help coping with the heterogeneity aspect, we also need an *ontology*, which allows us to describe in a uniform way the functionalities provided by each PEIS in the ecology and the data on which they operate, and to define the notion of compatibility used by the discovery mechanism.

Figure 3 illustrates our approach, in the case of a simple ecology consisting of a vacuum cleaner, a ceiling camera tracking system, and a workstation. Every PEIS is provided with a local directory of descriptions D and with a special component M that can access the descriptions and advertise them to the rest of the ecology. Some PEIS can be equipped with a special *configurator* component, denoted by $CONF$, that is capable of retrieving the descriptions and computing a meaningful configuration based upon the information stored in them. The configurator also takes care of deploying and monitoring the generated configuration. For the monitoring part, the configurator subscribes to `fail` signals from the connected PEIS, and re-triggers the configuration algorithm if any PEIS drops from the configuration for any reason. Note that not all PEIS need to include a configurator, and that multiple configurator components can exist in the ecology. Whenever a PEIS needs to generate a configuration to perform a task, it asks the service of an available configurator component.

The configurator component can be implemented using different approaches. In our project, we explore two complementary approaches for that. The first approach is a *plan-based*, centralized approach [11, 12]. In this approach, we use a global hierarchical planner to generate the (minimum cost) configuration for a given task. The second is a *reactive*, distributed approach [7]. In this approach, the configurator creates a local configuration, and assumes that the connected PEIS are able to recursively extend this configuration if needed. If they are not, the configurator receives a `fail` signal and tries a different local configuration. Both approaches provide some simple form of self-repair: if a PEIS signals that a functionality used in the current configuration is not available any more, the configurator tries to generate an alternative configuration.

The two approaches have the typical complementary strengths and weaknesses of plan-based and reactive approaches. The plan-based approach is guaranteed to find the optimal configuration if it exists, but it has problems

to scale up and it cannot easily cope with changes in the ecology. The reactive approach scales up smoothly and it can quickly adapt to changes in the state of the ecology, but it might generate non-optimal configurations and it might fail to find a configuration even if one exists. Eventually, we hope to be able to combine these two approaches into a hybrid configurator.

2.5 Experimental Validation

The PEIS-Ecology project follows a methodology which is strongly experimental: principles and techniques are systematically evaluated on physical platforms, and the results are used to refine these principles and techniques.

In order to follow this methodology, we have built a physical testbed facility called the PEIS-Home. This facility looks like a typical Swedish bachelor apartment (Figure 4). The PEIS-Home is equipped with a communication infrastructure and with a number of PEIS, including static cameras, mobile robots, multi-media devices, sensor nodes (motes), a refrigerator equipped with gas sensors and an RFID reader, and many more.

We have used this testbed to run a large number of experiments, with several aims: to test our technical solution and measure our progress; to validate the effectiveness of the developed techniques; to evaluate the acceptability of the PEIS-Ecology concept to human users, with special attention to elderly people; and to demonstrate the PEIS-Ecology concept to the society at large. Descriptions and videos of some of the experiments can be found on the PEIS-Ecology home page [14].

3 Future Directions

A PEIS-Ecology is meant to operate in the presence of, and in the service of, humans. It is therefore essential that the development of a PEIS-Ecology take into careful consideration the place of the humans in it.

The way in which a PEIS-Ecology interfaces with the human inhabitants is critical to its usability and acceptability. Humans should perceive the PEIS-Ecology either as one entity, or as a set of individual PEIS, depending on the context. In either case, they should use similar interaction modalities, and experience a natural interaction in compliance with social rules. The humans should also be made aware of what the PEIS-Ecology can afford to them, with special emphasis on those affordances which are most relevant given the current context.



FIGURE 4. *Two views of the PEIS-Home experimental testbed. Left: the kitchen, with a nose-equipped robot inspecting a smart fridge. Middle: the living room, with a human-interface robot in its parking position. Right: A Mote-IV mote monitoring temperature and humidity of a plant.*

In the reverse direction, a PEIS-Ecology should be able to incorporate humans among its parts, and to operate in symbiosis with them [5]. It should be able to infer the status and intentions of humans from observations, and adapt its behavior to that. For instance, if a human shows the intention to relax, the vacuum cleaner should move to a different room. A PEIS-Ecology should also be able to infer what the humans can afford to it: for instance, the vacuum cleaner could ask the human to empty its dust-bag if it knows that the human can afford that. Ideally, it should also be able to smoothly update its model of what a human user can afford to adapt to changes in this user, e.g., growing older.

Work more directly concerned with the inclusion of humans into a PEIS-Ecology has just started at the time of this writing, and it is expected to be a major drive of future developments in this area. Our approach is to see humans as just another species of PEIS in the ecology, which may use the (perceptual, acting, or information processing) functionalities provided by the rest of the PEIS-Ecology, and may provide functionalities to it. What makes humans a peculiar type of PEIS is that their goals and desires have a high priority status, and that they need to use dedicated human interface components to communicate with the rest of the ecology. Currently, we are exploring the use of template-based interface components to select, and make visible to the users, the capabilities of the PEIS-Ecology which are relevant in the current context [3]. For instance, when a human sits on the sofa after dinner, the options of bringing a drink, bringing the phone, or playing music, are made available to her. When the same human leaves

the house, the options of patrolling the house or keeping the house warm are offered instead.

4 Potential for Space Exploration

The concept of PEIS-Ecology was originally developed to allow the inclusion of robotic technologies in our homes, in order to improve the quality of life, safety and independence of citizens in general, and senior citizens in particular. Despite this, the PEIS-Ecology concept has several characteristic that makes it a promising approach to the inclusion of robot technologies in space exploration.

Thanks to its modularity, self-configurability and re-configurability, the PEIS-Ecology approach allows us to build robotic systems which are more flexible, adaptable, and robust than what can be achieved today. Flexibility, adaptability and robustness are clearly requirements for robots to be used in space exploration. Moreover, the PEIS-Ecology approach allows a distributed robot system to be assembled “on the fly” for a given task and context, by self-configuration of the PEIS which are available. An initial population of PEIS can be sent, say, on a planet for exploratory purposes, including a small number of powerful ones that can deploy a large number of simpler ones. New PEIS can be added later on, and participate in the formation of incrementally more powerful systems when needed. An interesting extension in this direction will be how to synchronize the hardware self-configuration of modular robots and the software self-configuration provided by the PEIS-Ecology framework.

Another future direction in the PEIS-Ecology concept which may have a strong potential for space exploration is the development of a full human-robot-environment symbiotic system, as outlined in the previous section. The integration of robots, sensors, augmented everyday objects, and humans as different species into one whole PEIS-Ecology would allow us to create a task-supporting ecosystem in the spaceship or on a far planet. This ecosystem would complement the life-supporting biological ecosystem, in providing physical and cognitive assistance to the astronaut in performing their everyday chores, as well as in dealing with exceptional situations. This would provide more time for the astronauts to perform their mission-related tasks. It would also provide increased safety, and offer the opportunity to engage the astronauts in entertaining and stress-releasing activities.

Here is a fully speculative attempt to put the above achievements on a time scale. By 2015, autonomous robotic technologies will start to be pervasively present in our homes, most probably in an “invisible” way, as well as most of today’s computers in our homes do not manifest themselves as such. By 2020, enough knowledge will have been gained on human-robot-environment symbiotic systems on earth, to allow their exploitation in the context of space missions. By 2030, the first robot ecology would be transported and deployed on another planet — to be later expanded by sending additional robots and devices.

5 Conclusions

The idea to integrate robots and smart environments is starting to pop up at several places and under several names, including network robot systems [13], intelligent space [9], sensor-actuator networks [6], ubiquitous robotics [8], artificial ecosystems [16], and still others. A few projects were recently started with the aim to explore the scientific, technological and practical implications of this integration. Currently the largest efforts are probably the Network Robot Forum [13], the U-RT project at AIST [10], and the Korean Ubiquitous Robot Companion program [8]. The PEIS-Ecology project presented in this paper is part of the latter effort. This project is distinct in its emphasis on the study of the fundamental scientific principles that underlie the design and operation of an ubiquitous robotic system.

The PEIS-Ecology project has been active for 3.5 years at the time of this writing. A number of scientific and technological achievements have already been obtained

in this context, including the three ones reported in this note: the development of an open-source middleware, the study of the integration between physical and digital interaction with the world, and new techniques for self-reconfiguration and re-configuration. Source code, videos and scientific papers relative to these achievements are available at the PEIS-Ecology web site [14]. The next important step in this development will be the inclusion of humans into a PEIS-Ecology. We expect that the PEIS-Ecology concept may become relevant to space missions at least in two ways: as a way to deploy robotic missions for planetary exploration, and as a way to provide assistance to astronauts.

Acknowledgments

The vision and the achievements summarized above are the results of the cooperation of many people who are involved in the PEIS-Ecology project. These include Mathias Broxvall, Marcello Cirillo, Marco Gritti, Donatella Guarino, Lars Karlsson, Kevin LeBlanc, Amy Loutfi, Robert Lundh, Federico Pecora, Jayedur Rashid, and Beom-Su Seo. Financial support for the PEIS-Ecology project was provided by ETRI (Electronics and Telecommunications Research Institute, Korea) through the project “Embedded Component Technology and Standardization for URC (2004-2008)”. Additional support was provided by Vetenskapsrådet (the Swedish Research Council), by CUGS (the Swedish Graduate School in Computer Science), and by Örebro University. The author wishes to thank the Advanced Concepts Teams of ESA for offering the opportunity of a most enjoyable scientific exchange.

References

- [1] M. Broxvall, S. Coradeschi, A. Loutfi, and A. Saffiotti. An ecological approach to odour recognition in intelligent environments. In *Proc of the IEEE Int Conf on Robotics and Automation*, pages 2066–2071, Orlando, FL, 2006.
- [2] M. Broxvall, M. Gritti, A. Saffiotti, B. Seo, and Y. Cho. PEIS ecology: Integrating robots into smart environments. In *Proc of the IEEE Int Conf on Robotics and Automation*, pages 212–218, Orlando, FL, 2006.
- [3] M. Broxvall, A. Loutfi, and A. Saffiotti. Interacting with a PEIS ecology. In *Proc of the IEEE Int*

- Symp on Robot and Human Interactive Communication (Ro-Man)*, 2007.
- [4] S. Coradeschi and A. Saffiotti. An introduction to the anchoring problem. *Robotics and Autonomous Systems*, 43(2-3):85-96, 2003.
- [5] S. Coradeschi and A. Saffiotti. Symbiotic robotic systems: Humans, robots, and smart environments. *IEEE Intelligent Systems*, 21(3):82-84, 2006.
- [6] F. Dressler. Self-organization in autonomous sensor/actuator networks. In *Proc of the 19th IEEE Int Conf on Architecture of Computing Systems*, 2006.
- [7] M. Gritti, M. Broxvall, and A. Saffiotti. Reactive self-configuration of an ecology of robots. In *Proc of the ICRA Workshop on Networked Robot Systems*, Rome, Italy, 2007.
- [8] J. Kim, Y. Kim, and K. Lee. The third generation of robotics: Ubiquitous robot. In *Proc of the 2nd Int Conf on Autonomous Robots and Agents*, Palmerston North, New Zealand, 2004.
- [9] J. Lee and H. Hashimoto. Intelligent space — concept and contents. *Advanced Robotics*, 16(3):265-280, 2002.
- [10] O. Lemaire, K. Ohba, and S. Hirai. Dynamic integration of ubiquitous robotic systems using ontologies and the rt middleware. In *Proc of the 3rd Int Conf on Ubiquitous Robots and Ambient Intelligence*, Seoul, Korea, 2006.
- [11] R. Lundh, L. Karlsson, and A. Saffiotti. Plan-based configuration of an ecology of robots. In *Proc of the IEEE Int Conf on Robotics and Automation*, Rome, Italy, 2007.
- [12] R. Lundh, L. Karlsson, and A. Saffiotti. Autonomous functional configuration of a network robot system. *Robotics and Autonomous Systems*, 2008. To appear.
- [13] Network robot forum home page. <http://www.scat.or.jp/nrf/English/>, 2008.
- [14] The PEIS ecology project. <http://www.aass.oru.se/~peis/>, 2008.
- [15] A. Saffiotti and M. Broxvall. PEIS ecologies: Ambient intelligence meets autonomous robotics. In *Proc of the Int Conf on Smart Objects and Ambient Intelligence (sOc-EUSAI)*, pages 275-280, Grenoble, France, 2005.
- [16] A. Sgorbissa and R. Zaccaria. The artificial ecosystem: a distributed approach to service robotics. In *Proc of the IEEE Int Conf on Robotics and Automation*, pages 3531-3536, 2004.



Acta Futura 3 (2009) 43-49

DOI: 10.2420/AF03.2008.36

**Acta
Futura**

Integration of Cellular Biological Structures Into Robotic Systems

J. GOUGH, G. JONES, C. LOVELL, P. MACEY,
H. MORGAN, F. D. REVILLA, R. SPANTON, S. TSUDA, AND K.-P. ZAUNER*

*School of Electronics and Computer Science, University of Southampton,
Southampton SO17 1BJ, United Kingdom.*

1 Introduction

Ideas from biological information processing have influenced numerous robot control architectures. Most of them draw their inspiration from neural networks. Many cognitive capabilities, however, can only be understood from a system level perspective that integrates body, control, and environment. This insight led to robot designs that exploit the physical characteristics of a robot's body [14]. If one takes the direction of these investigations further, one arrives at the point where control is not only exploiting the physics of the body, but is in itself directly driven by physics.

Arguably this step is necessary to narrow the performance gap between man-made devices and biological systems. All organisms require information processing to defend their intricate organisation against the onslaught of physical entropy. As a consequence organisms exhibit an intriguing sophistication in overcoming computationally difficult challenges. This is the case even for simple life forms and individual cells. Robots and organisms both need to process sensory information from a range of modalities and make decisions based on noisy ambiguous data. Not only do they need to process the

sensory inputs and respond in real-time with appropriate actions, but they need to do this under the resource restrictions imposed by their size. Organisms approach this challenge by directly exploiting the physico-chemical properties of their molecular materials. This mode of information processing stands in stark contrast to present computing technology where the formalisms of which are defined in disregard of the physical substrate employed to implement them. Brooks observed that "the matter that makes up living systems obeys the laws of physics in ways that are expensive to simulate computationally" [4] and Conrad argued that programmable computing is necessarily inefficient compared to nature's information processing [5, 6]. From what has been stated above, it appears worthwhile to give more attention to the role computational substrates may play in cognitive systems [19]. Taking heed of the clues from biology, macromolecular materials appear particularly attractive [23].

2 On-board Cellular Robot Control

We endeavour to sketch here a potential route to recruiting some of nature's information processing efficiency for technical applications: the integration of living cells

*Corresponding author. E-mail address: kpz@ecs.soton.ac.uk



FIGURE 1. A tethered hexapod robot optically interfaced with a *Physarum polycephalum* cell [21].

into bio-electronic hybrid robot control devices. Living cells provide the complete infrastructure one would like to have in place for molecular information processing. They also add complexity and new challenges, but it is worthwhile to explore whether these obstacles can be surmounted.

In the approach presented, selected features of the robot's environment are mapped into features of the environment of a living cell. Conversely, the response of the cell is mapped back onto the actuators of the robot. Fig. 1 shows an early implementation of this concept. The cell employed in the experiments is the plasmodial stage of the slime mould *Physarum polycephalum*. This large multinucleated cell can be found on decaying wood and feeds on bacteria and organic matter. Despite its macroscopic size (it typically grows to $>100\text{ cm}^2$) the cell acts as a single integrated organism. Too large to rely on diffusion, it uses rhythmic contractions to distribute materials within its cell body [7]. These oscillations of the protoplasm are due to hydrostatic pressure changes driven by actomyosin. The local frequency of the oscillations is affected by chemical and physical stimuli [9].

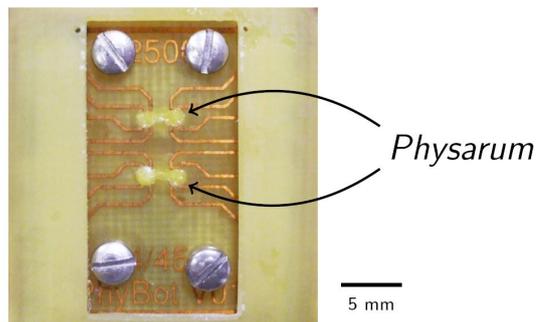


FIGURE 2. Electrical interfacing of the *Physarum* plasmodium.

Central to our approach is the cell-robot interface. Both, optical and electronic interfaces for plasmodia of the slime mold have been implemented. Our first approach to cellular robot control had purposely set aside the question of whether it is practicable to integrate a cell-based controller into a small robot [20]. In recent work we have addressed this issue. The current architecture uses the oscillator design suggested by Takamatsu and Fujii [17] but with a printed circuit board (PCB) applied as negative-mask.

The mechanical oscillations of the plasmodium that were previously detected as local changes in light transmission of the cell body, can also be detected by the change in impedance among the electrodes patterned on the PCB [8, 15].

Fig. 2 shows a device with two independent plasmodia (labeled *Physarum* in the figure). Each is grown into an opening in the circuit board that confines the plasmodium to dumbbell shape with a left and a right well connected by a channel. Two pairs of electrodes are available for each well. The image shows the underside of the chip through an aperture in the circuit board carrying the circuit for measuring the impedance. During operation, this face of the PCB would be covered with a layer of agar. To monitor the oscillations of the slime mould cell, we employ an integrated circuit capable of measuring impedance at up to 100 kHz. For further details refer to the Methods section.

The impedance across each of the eight pairs of electrodes is measured once per second, the maximal rate our multiplexed setup permits. Fig. 3 shows the magni-

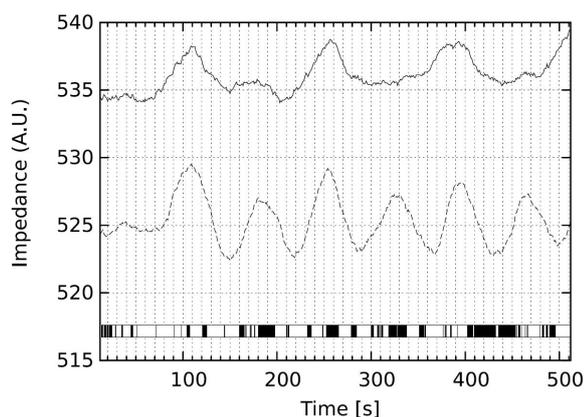


FIGURE 3. Change in the magnitude of the impedance at 100 kHz concomitant with the volume oscillations of two parts of a single slime mould cell.

tude of the impedance signal originating from two wells, i.e., two ends of a single cell, after a moving average filter has been applied to reduce noise. The two curves show the moving averages over 15 samples, recorded from the right and left well of the plasmodium. The phase-relationship between the two wells is shown in the bar across the bottom of the figure where white and black indicate in- and out-of-phase, respectively. In the experiment described here, this binary phase-relation is used to steer a robot. The minimalist robot, pictured in Fig. 4, was inspired by Braitenberg's vehicles [3] and is designed as a test platform for unconventional controllers.

In the present experiments a simple one bit actuator control modeled after the control of actuation in bacterial chemotaxis [1, 16] has been employed. The control bit toggles between the state of turning in a random direction or moving forward. The state is updated at a rate of 1 Hz, whereby in-phase oscillation of the two wells is mapped to random turning and out-of phase oscillation is mapped to straight moves. A typical segment of a trajectory without stimulation is shown in Fig. 5. An illuminated target on the robot is tracked for position and orientation with an overhead camera. If the robot reaches the boundary of the arena, forward commands are ignored and only turns are executed. So far this prototype only implements the actuator side; the facility to transmit the sensory stimulation to the slime mould is not yet completed. Nevertheless, preliminary tests indi-

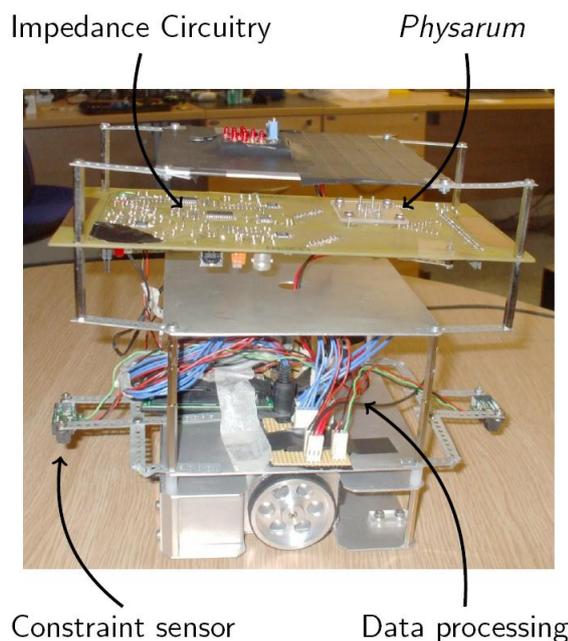


FIGURE 4. Robotic platform with the *Physarum* chip installed on the impedance circuit board.

cate that in principle the integration of living cells into a robot is feasible.

3 Vision: Living Devices

Microbial cells are routinely employed in many engineered processes. Their ability to self-reproduce offers a cheap and fast route to deploy complex nano-scale systems. Parallel to the progress in molecular biology, the traditional breeding (e.g., to optimise baker's yeast) has been augmented with direct modification of cell lines to introduce new capabilities (e.g., production of human insulin). What started as metabolic engineering with static changes to the genome of the cells, is now, as "synthetic biology," on the way to engineer dynamic control structures and may eventually allow the design of application specific cells [2]. In addition to self-reproduction, cells provide quality control and testing at the point of assembly for many of their macromolecular products and can replenish their components and thus operate under adverse conditions.

Another important development is the increasing sophistication of lab-on-chip and microfluidic technolo-

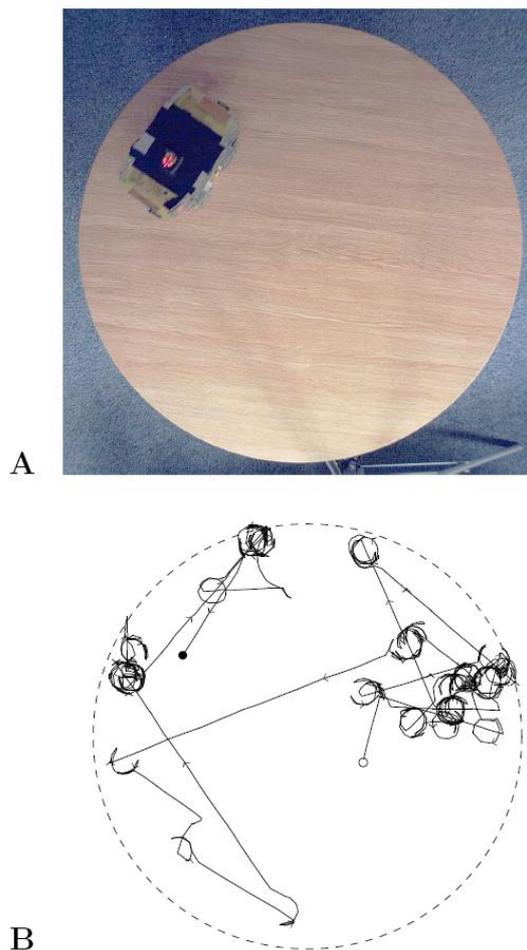


FIGURE 5. The arena to which the robot is confined in the view of the tracking camera (A), and the trajectory of the robot corresponding to the signals shown in Fig. 3; \circ marks the start and \bullet the end of the segment (B).

gies that make it possible to maintain a controlled microenvironment for a cell. The confluence of both of these lines of research opens up a new engineering direction where living cells are tightly integrated with conventional technology. In the resulting devices, living cells are an essential component. Examples are “living sensors” [13, 22] and “living pumps” [18, 11]. In the coming decades such an approach will almost certainly be explored for elementary cognitive systems and we see the ongoing work described in the previous section as a small step in this direction.

4 Potential for Space Applications

The key features of cells that may be of potential interest in the space field are:

1. Self-repair capability
2. Self-reconfiguration capability
3. Efficient in material requirement
4. Efficient in energy usage

Cells, in general, will recognise a wide variety of damages and take steps to repair themselves. Some extremophile microbes, in particular, possess elaborate innate mechanisms to actively maintain their molecular organisation. The material efficiency of cells usually permits redundancy of functional components (especially proteins) and information carriers (double stranded DNA) and thus enables them to autonomously recover from many injuries. A slime mould plasmodium can contain many millions of redundant copies of its genetic information within a single cell. Potentially cells could offer a unique solution to the problem of transporting a complex functional system over a long period through a harsh environment when shielding is impractical due to weight constraints. It is also possible to imagine applications of cells to grow a functionalised surface such as a sensory-skin for a robot. Only a small volume of seed cells (or spores) would need to be protected from radiation damage during transit and after formation such a surface may be able to self-heal when damaged.

Of course there are also problems. Devices containing living cells pose the danger of spreading organisms from Earth or may interfere with the search for indicators of non-terrestrial life. Cells do have different requirements from inanimate materials, have a quite restricted operating temperature range, need to be protected against loss of moisture, and in general have idiosyncratic supply needs.

5 Conclusion

Robots and organisms face the same problem: they need to act in real-time in a complex environment and are severely restricted in material and energy that can be allocated to their information processing needs. Organisms, however, spectacularly outperform current robots. It appears increasingly likely that the choice of computational substrate is critical to their success.

The integration of living cells into robotic systems is feasible, albeit at its infancy. Cells offer many properties that are unlikely to come within reach of conventional technologies within the foreseeable future, in particular self-reproduction and self-repair of complex heterogeneous structures. Devising bio-hybrid architectures in which cells confer their unique characteristics in a beneficial manner to the system as a whole, will be among the key challenges in this field.

Conceivable early application domains in space research may lie in non-critical subsystems where the integration of living devices potentially offers the opportunity to deploy systems of high complexity with a reasonable chance to recover, or grow functionality after extended periods of exposure to high electromagnetic radiation flux. It goes without saying that any application scenario needs to exclude the possibility of contaminating celestial bodies.

A Methods

Electrodes were patterned on a small circuit board (PCB); the wells drilled (cf., Fig. 2) and a channel cut between each pair of wells. The design of the electrode pattern allows for two independent plasmodia. Each plasmodium is in a dumbbell configuration with two 1.6 mm diameter wells at a centre-distance of 2.5 mm and connected by a 0.4 mm wide channel. Two pairs of electrodes are available on each well for a total of 16 electrodes on the board. The copper side of the circuit board is laminated except at the areas where it connects to the impedance circuit. To provide the plasmodium with a moist surface the board is placed with the laminated side facing a layer of agar gel.

Physarum polycephalum was cultured on 1.5% agar and fed with oat flakes. It is possible to let the slime mould grow onto the agar surface masked with the circuit board. However, to accelerate the preparation for the experiments we proceed as follows. First a tubular section of a plasmodium is excised from a large culture and implanted into the channel. Next the wells, i.e., the holes drilled in the circuit board are filled with material from the anterior, fan-like parts of the culture. By covering the top side of the circuit board with a thin, gas-permeable PDMS membrane the plasmodium is fully enclosed, which enables experiments of several hours duration. A perspex frame holds the layers together. We refer to this stack of agar sheet, laminate, (copper layer) PCB and PDMS as a “chip”. After an incubation

period of 2–3 h in the dark, during which the material in the dumbbell shaped cut-out fuses into a single plasmodial cell, the chip is ready for experiments. We typically prepare several plasmodia in chips and select those that show a strong oscillation signal.

A custom circuit based on an impedance converter/network analyser (AD5933, Analog Devices, www.analog.com) and two multiplexers (ADG732, Analog Devices) in combination with custom software enables the monitoring of the cellular oscillations through either a universal serial bus or I²C interface [12]. The actual implementation of the circuit used in the experiments described here, is a prototype with easy access to all components and the ability to isolate blocks for testing. We expect to reduce the layout to less than a quarter of its current size (cf. Fig. 4). Impedance is measured at 100 kHz applying an excitation of ≈ 1 V peak-to-peak.

The impedance circuit is connected through an I²C bus to a gumstix computer (www.gumstix.com), which samples all eight electrode pairs at ≈ 1 Hz and stores the information in flash memory. The data from all electrodes on both plasmodia on the chip is recorded for later analysis. For driving the robot, however, one of the plasmodia is selected. For each of the two wells filled by this plasmodium, only the data from the pair of electrodes that show the stronger oscillation is selected for further processing. To reduce noise the moving average over 15 samples is computed (curves in Fig. 3). This provides a sufficient signal even though not much provision for screening against noise was made in the prototype, except separating the power supply for the impedance circuitry from the supply of the robot and gumstix computer. After the noise filtering, the drifting-component in the signal is removed by subtracting the 15 s delayed signal. Then the phase-relationship between the two signals originating from the left and right well is classified according to whether both signals have equal or opposite sign (bar at the bottom of Fig. 3).

The impedance circuit board carrying the *Physarum* chip and the gumstix are mounted on custom designed wheeled robot base [10]. A simple rule is used to map the phase-relationship between the two wells onto the motion of the robot: If the signs of the signals are equal the robot takes a random turn, if the signs differ the robot moves straight forward. Within the 1 second period between state updates the robot can move at most 9 cm. The robot is confined to a round table with 1 m diameter by means of an infrared sensor. If the sensor moves off the table, further “forward” commands are ig-

nored by the low-level driver circuit of the robot until a rotation brings this constraint sensor back onto the table surface. For tracking of the robot a network camera (AXIS 206M, www.axis.com) is mounted above the table and images captured at a rate of 3 frames per second. An illuminated arrow-shaped target mounted on the robot is located automatically in the frames to obtain the trajectory of the robot.

Acknowledgements

This research was funded in part by the University of Southampton Life Sciences Interface Initiative, supporting FDR and ST. Funds from Microsoft Research Cambridge, The Nuffield Foundation and the Southampton Nano Forum provided summer internships for JG, PM, and RS.

References

- [1] J. Adler and W.-W. Tso. "Decision"-making in bacteria: Chemotactic response of *Escherichia coli* to conflicting stimuli. *Science*, 184:1292–1294, 1974.
- [2] W. J. Blake and F. J. Isaacs. Synthetic biology evolves. *Trends in Biotechnology*, 22(7):321–324, 2004.
- [3] V. Braitenberg. *Experiments in Synthetic Psychology*. MIT Press, Cambridge, MA, 1984.
- [4] R. A. Brooks. The relationship between matter and life. *Nature*, 409:409–411, 2001.
- [5] M. Conrad. The price of programmability. In R. Herken, editor, *The Universal Turing Machine: A Fifty Year Survey*, pages 285–307. Oxford University Press, New York, 1988.
- [6] M. Conrad. Scaling of efficiency in programmable and non-programmable systems. *BioSystems*, 35:161–166, 1995.
- [7] A. C. H. Durham and E. B. Ridgway. Control of chemotaxis in *Physarum polycephalum*. *The Journal of Cell Biology*, 69:218–223, 1976.
- [8] R. Halvorsrud, I. Giaever, and M. M. Laane. Patterns of oscillation during mitosis in plasmodia of *Physarum polycephalum*. *Protoplasma*, 188:12–21, 1995.
- [9] Z. Hejnowicz and K. E. Wohlfarth-Bottermann. Propagated waves induced by gradients of physiological factors within plasmodia of *Physarum polycephalum*. *Planta*, 150:144–152, 1980.
- [10] G. Jones. Robotic platform for molecular controlled robots. Part III Project Report, School of Electronics and Computer Science, University of Southampton, May 2006.
- [11] M. J. Kim and K. S. Breuer. Controlled mixing in microfluidic systems using bacterial chemotaxis. *Anal. Chem.*, 79:955–959, 2007.
- [12] P. Macey. Impedance spectroscopy based interfacing with a living cell for biosensors and bioprocessors. Part III Project Report, School of Electronics and Computer Science, University of Southampton, May 2007.
- [13] D. E. Nivens, T. E. McKnight, S. A. Moser, S. J. Osbourn, M. L. Simpson, and G. S. Saylery. Bioluminescent bioreporter integrated circuits: potentially small, rugged and inexpensive whole-cell biosensors for remote environmental monitoring. *Journal of Applied Microbiology*, 96:33–46, 2004. doi:10.1046/j.1365-2672.2003.02114.x.
- [14] R. Pfeiffer and C. Scheier. *Understanding Intelligence*. The MIT Press, Cambridge, MA, 1999.
- [15] F. D. Revilla, K.-P. Zauner, and H. Morgan. *Physarum polycephalum* on a chip. In J.-L. Viovy, P. Tabeling, S. Descroix, and L. Malaquin, editors, *The Proceedings of μ TAS 2007*, volume 2, pages 1089–1091, 2007.
- [16] B. E. Scharf, K. A. Fahrner, L. Turner, and H. C. Berg. Control of direction of flagellar rotation in bacterial chemotaxis. *Proc. Natl. Acad. Sci. USA*, 95:201–206, 1998.
- [17] A. Takamatsu and T. Fujii. Time delay effect in a living coupled oscillator system with the plasmodium of *Physarum polycephalum*. *Physical Review Letters*, 85:2026–2029, 2000.
- [18] Y. Tanaka, K. Sato, T. Shimizu, M. Yamato, T. Okanobd, and T. Kitamori. A micro-spherical heart pump powered by cultured cardiomyocytes. *Lab on a Chip*, 7:207–212, 2007.

- [19] S. Tsuda, K.-P. Zauner, and Y.-P. Gunji. Computing substrates and life. In S. Artmann and P. Dittrich, editors, *Explorations in the Complexity of Possible Life: Abstracting and Synthesizing the Principles of Living Systems*, Proceedings of the 7th German Workshop on Artificial Life, pages 39–49, Jena, Germany, July 26–28 2006. IOS Press.
- [20] S. Tsuda, K.-P. Zauner, and Y.-P. Gunji. Robot control: From silicon circuitry to cells. In A. J.Ijspeert, T. Masuzawa, and S. Kusumoto, editors, *Biologically Inspired Approaches to Advanced Information Technology (BioADIT'06)*, Osaka, Japan, January 26–27, 2006, *Proceedings*, volume 3853 of *LNCIS*, pages 20–32, Berlin, 2006. Springer.
- [21] S. Tsuda, K.-P. Zauner, and Y.-P. Gunji. Robot control with biological cells. *BioSystems*, 87:215–223, 2007.
- [22] R. Vijayaraghavan, S. K. Islam, M. Zhang, S. Ripp, S. Caylor, N. D. Bull, S. Moser, S. C. Terry, B. J. Blalock, and G. Sayler. A bioreporter bioluminescent integrated circuit for very low-level chemical sensing in both gas and liquid environments. *Sensors and Actuators B*, 123:922–928, 2007.
- [23] K.-P. Zauner. Molecular information technology. *Critical Reviews in Solid State and Material Sciences*, 30(1):33–69, 2005.

Panel 2: “*Vis Viva*” from Space

LEIBNITZ first defined Energy as “*vis viva*”, “living force”. Energy (from the Greek ενεργός = “at work”, “active”) is the key ingredient to allow living species to act on the environment for satisfying fundamental needs. At the same time energy is “shaped” by the environment in its different forms while obeying to conservation laws. So one can see energy as the trace of an active (“living”) environment. There is no doubt humans have been gifted on Earth with a particularly “active” environment from the energetic point of view. Both renewable and expendable sources of energy have been abundant on our planet. In the course of the history human civilisation has learnt how to harness some of these resources, mostly nonrenewable, to speed up its evolution. Yet other valuable resources have been largely ignored.

It is only due to environmental concerns of the last few years that one important fact was finally highlighted: we have still not fully learned how to make the best use out of many different energy resources of our “living” environment. It is hoped that the call towards research on renewable and sustainable energy production will make us evolve in this direction.

In space, at least as far as our solar system, energy conditions seem not to be as favorable as on Earth. In the interplanetary medium, for example, the only apparent large scale power source is the Sun radiation, harnessed since the very first satellites were put into orbit. Energy in space is so valuable, that space engineers are arguably the most energy-conscious engineers of all industries. Not only spacecraft are constructed with as high an energy efficiency as possible, contrary to Earth-based systems space energy systems are inherently holistic, including the entire chain up to eventual waste products (e.g. heat) into the design.

But can we really label the space environment as energy-starved? Maybe not, or at least not always. On one hand space can offer conditions which resemble the terrestrial ones. Thermal gradients, winds, seismic phenomena can possibly be found and exploited for power production using techniques adapted or inspired by the ones used on Earth. Bio-waste energy conversion, a promising power production scheme on Earth, may one day become a standard procedure for human space ex-

ploration missions. On the other hand some unique conditions typical of the space environment (e.g. vacuum, free fall conditions, presence of plasma) may allow for alternative ways of power extraction which are not possible or viable on Earth. Two of these innovative schemes are presented in this panel.

How far are we from being able to fully exploit space as a living energy environment? The goal of this panel will be to address this and other important questions related to the space energy field and its future.

Claudio Bombardelli
Panel coordinator



Kinetic Energy Harvesting

S. P. BEEBY*, R. N. TORAH, M. J. TUDOR

School of Electronics and Computer Science, University of Southampton, UK

Abstract. This paper reviews kinetic energy harvesting as a potential localised power supply for wireless applications. Harvesting devices are typically implemented as resonant devices of which the power output depends upon the size of the inertial mass, the frequency and amplitude of the driving vibrations, the maximum available mass displacement and the damping. Three transduction mechanisms are currently primarily employed to convert mechanical into electrical energy: electromagnetic, piezoelectric and electrostatic. Piezoelectric and electrostatic mechanisms are best suited to small size MEMS implementations, but the power output from such devices is at present limited to a few microwatts. An electromagnetic generator implemented with discrete components has produced a power 120 W with the highest recorded efficiency to date of 51% for a device of this size reported to date. The packaged device is 0.8 cm^3 and weighs 1.6 grams. The suitability of the technology in space applications will be determined by the nature of the available kinetic energy and the required level of output power. A radioactively coupled device may present an opportunity where suitable vibrations do not exist.

1 Introduction

Wireless systems are becoming ubiquitous and the use of wireless devices offers several advantages over existing, wired methodologies. These include flexibility, ease of implementation and the ability to place sensors in previously inaccessible locations. Furthermore, in the case of wired solutions, the wires (and associated connectors) are often a source of failure, add weight and present a considerable cost. The ability to retrofit systems without having to consider issues such as cabling, offers a significant advantage in applications for areas like condition-based monitoring [5].

At present, many wireless sensor nodes are battery-powered and operate on an extremely economical energy budget since continuous battery replacement is not an option for networks with thousands of physically embedded nodes [21]. The low power characteristics of wireless sensor network components and the design of the system architecture are crucial to the longevity of the sensor nodes. Low power wireless communications protocols (e.g. IEEE 802.15.4 Standard and Zigbee [10]) and the use of intelligence at the sensor node to perform signal processing on the raw sensor data, execute communications protocols and manage the node's power consumption have resulted in very low-power wireless sensor nodes. These provide the opportunity for alternative types of power source to traditional batteries. Renewable power can be obtained by generating electrical

*Corresponding author. E-mail address: spb@ecs.soton.ac.uk

energy from light, thermal and kinetic energy present within the sensor's environment. These sources can be used as either a direct replacement or to augment the battery, thereby increasing the lifetime and capability of the network [14, 12, 1, 13] and mitigate the environmental impact caused by issues surrounding the disposal of batteries.

The subject of this paper is kinetic energy generators, which convert energy in the form of movement present in the application environment into electrical energy [4]. Kinetic energy is typically present in the form of vibrations, random displacements or forces and is typically converted into electrical energy using electromagnetic, piezoelectric or electrostatic mechanisms. Suitable vibrations can be found in numerous applications including common household goods (e.g. fridges, washing machines, microwave ovens), industrial plant equipment, automobiles and aeroplanes and structures such as buildings and bridges [23]. Human-based applications are characterised by low frequency high amplitude displacements [24, 27].

The amount of energy generated by this approach depends fundamentally upon the quantity and form of the kinetic energy available in the application environment and the efficiency of the generator and the power conversion electronics. The following sections will discuss the fundamentals of kinetic energy harvesting and the different transduction mechanisms that may be employed. These mechanisms will then be illustrated by the results from a recently completed EU project Vibration Energy Scavenging (VIBES), the results from which represent the state of the art.

2 Theory

Kinetic energy harvesting requires a transduction mechanism to generate electrical energy from motion and a mechanical system that couples environmental displacements to the transduction mechanism. The design of the mechanical system should maximise the coupling between the kinetic energy source and the transduction mechanism and will depend entirely upon the characteristics of the environmental motion. Vibration energy is best suited to inertial generators with the mechanical component attached to an inertial frame which acts as the fixed reference (see Figure 1).

The inertial frame transmits the vibrations to a suspended inertial mass producing a relative displacement between them. Such a system will possess a resonant

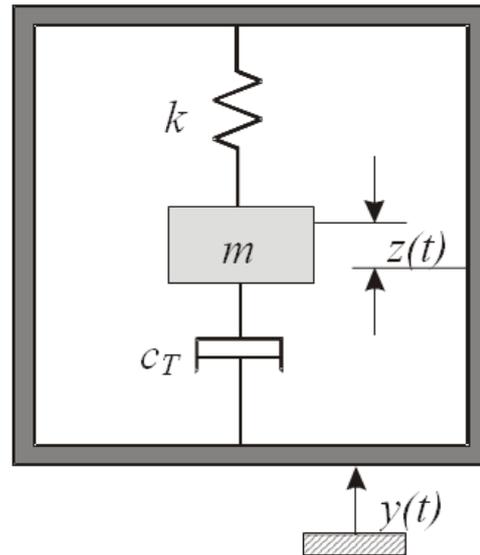


FIGURE 1. Inertial mass with spring stiffness k , mass displacement $z(t)$, damping coefficient c_T and input amplitude $y(t)$

frequency which must be designed to match the characteristic frequency of the application environment. The resonant approach magnifies the environmental vibration amplitude by the quality factor (Q -factor, Q_T) of the resonant system. The Q -factor is limited by damping effects such as unwanted losses and the transduction mechanism used to convert kinetic into electrical energy.

The limitation of the resonant approach is the need to operate at a particular frequency. This limitation can be overcome in part by using a wide bandwidth device, but this approach will reduce the peak power generated. Alternatively, bi-stable structures that flip between states can be employed. This approach will work at any frequency provided the external movement provides sufficient acceleration, but is complicated to achieve in practice and will result in reduced power densities in the majority of applications, depending upon the nature of the excitation [20, 9].

The theory of inertial-based generators is well documented [25] and will only be briefly summarized here. Assuming the generator is driven by a harmonic base excitation $y(t) = Y \sin(\omega t)$, it will move out of phase with the mass at resonance resulting in a net displacement, $z(t)$, between the mass and the frame. The aver-

age power dissipated within the damper (i.e. the power extracted by the transduction mechanism and the power lost through parasitic damping mechanisms) is given by:

$$P_{av} = \frac{m\xi_T Y^2 \left(\frac{\omega}{\omega_n}\right)^3 \omega^3}{\left[1 - \left(\frac{\omega}{\omega_n}\right)^2\right]^2 + \left[2\xi_T \left(\frac{\omega}{\omega_n}\right)\right]^2} \quad (1)$$

where ξ is the total damping ratio given by $\xi_T = c_T/2m\omega_n$. Since this equation is valid for steady state conditions, P_{av} is equal to the kinetic energy supplied per second by the application vibrations. Maximum power dissipation within the generator occurs when the device is operated at ω_n and in this case the P_{av} is given by:

$$P_{av} = \frac{mY^2\omega_n^3}{4\xi_T} \quad (2)$$

Equation (2) suggests the following rules:

- power varies linearly with the mass;
- power increases with the cube of the frequency; and
- power increases with the square of the base amplitude.

Rules (b) and (c) are dependant upon the base excitation i.e. the accelerations present in the application environment. Since the peak acceleration of the base, A , is given by $A = \omega^2 Y$ and damping factor is related the damping ratio by $c_T = 2m\omega_n \xi_T$, equation (2) can also be written in the form:

$$P_{av} = \frac{(mA)^2}{2c_T} \quad (3)$$

These equations emphasize the need to understand the vibrations present in the intended application when designing an inertial generator. However, one cannot simply choose a particular frequency of operation based upon the power output alone. The inertial mass displacement will be limited to a given finite value, z_{max} , depending upon the size of the generator, its design and material limitations. This is especially relevant in the case of MEMS generators. Furthermore, z_{max} will be a multiple Q_T times larger than Y where Q_T is given by equation (4):

$$Q_T = \frac{\omega_n m}{c_T} = \frac{1}{2\xi_T} \quad (4)$$

The relationship between Q_T and the electrical and parasitic damping factors is given by equation (5) where Q_{OC} is the open circuit Q -factor, i.e. $1/2\xi_P$, and Q_E is equal to $1/2\xi_E$.

$$\frac{1}{Q_T} = \frac{1}{Q_{OC}} + \frac{1}{Q_E} \quad (5)$$

Taking z_{max} into consideration, average power can also be expressed as:

$$P_{av} = \frac{m\omega_n^3 Y z_{max}}{2} \quad (6)$$

Incorporating the parasitic and electrical damping into equation (2) gives the average power delivered to the electrical domain:

$$P_{avlec} = \frac{m\xi_E Y^2 \omega_n^3}{4(\xi_P + \xi_E)^2} \quad (7)$$

Maximum power is delivered to electrical domain when $\xi_E = \xi_P$ i.e. damping arising from the electrical domain should equal mechanical losses. This applies when $z(t)$ is less than or equal to z_{max} . In this case equation (7) simplifies to:

$$P_{av} = \frac{mY^2\omega_n^3}{16\xi_P} \quad (8)$$

In many cases and especially for small devices, it is best to maximise the electrical damping and minimise parasitic losses due to the restricted z_{max} for a particular device.

3 Transduction

The transduction mechanism itself can generate electricity by exploiting the mechanical strain or relative displacement occurring within the system. The strain effect exploits a deformation and typically employs active materials (e.g. piezoelectric) to generate electrical energy from the resulting strain. In the case of relative displacement, either the velocity or position can be used. Velocity is typically associated with electromagnetic transduction whilst relative position is associated with electrostatic transduction. Each transduction mechanism exhibits different damping characteristics and this should be taken into consideration when modelling the generators.

Piezoelectric Transduction The piezoelectric effect arises when certain materials subject to mechanical strain become electrically polarized. Conversely, these materials deform when exposed to an electric field. Piezoelectric materials are widely available in many forms including single crystal (e.g. quartz), piezoceramic (e.g. lead zirconate titanate or PZT), thin film

(e.g. sputtered zinc oxide), screen printable thick-films based upon piezoceramic powders [2] and polymeric materials such as Polyvinylidene fluoride (PVDF) [17].

Piezoelectric materials exhibit anisotropic characteristics i.e. the properties of the material differ depending upon the direction of forces and orientation of the polarisation and electrodes. The level of piezoelectric activity of a material is defined by a series of constants used in conjunction with the axes notation. The piezoelectric strain constant, d , relates the strain developed to the applied field. Piezoelectric generators that rely on a compressive strain applied perpendicular to the electrodes exploit the d_{33} coefficient of the material. Typically, in the case of piezoelectric films or piezoelectric elements bonded onto substrates, the elements are coupled with a transverse strain parallel to the electrodes and utilise the d_{31} coefficient. Such an arrangement provides mechanical amplification of the applied stresses.

Another important constant determining the generation of electrical power is the electro-mechanical coupling coefficient, k . This describes the efficiency with which the energy is converted by the material between electrical and mechanical domains. It is given by equation (9), where W_e and W_m are electrical and mechanical energy respectively.

$$k^2 = \frac{W_e}{W_m} \quad (9)$$

Piezoelectric properties vary with age, stress and temperature. The aging rate of a piezoceramic is dependant on the construction methods and the material type. The changes in the material tend to be logarithmic with time, thus the material properties stabilise with age. The aging process is accelerated by the stress applied to the material and this should be considered in cyclically loaded energy harvesting applications. Soft piezoceramic compositions, such as PZT-5H, are more susceptible to stress induced changes than the harder compositions such as PZT-5A. Maximum temperature of a piezoceramic is limited to the Curie point above which it becomes de-poled. The application of stress can lower this Curie temperature making high stress, high temperature applications challenging. The piezoelectric constants for common materials, soft and hard lead zirconate titanate piezoceramics (PZT-5H and PZT-5A), barium titanate (BaTiO₃) and PVDF, are given in table 1.

Electromagnetic Transduction Electromagnetic induction is the generation of electric current in a conductor located within a magnetic field. The conductor

Property	PZT-5H	PZT-5A	BaTiO ₃	PVDF
d_{33} (10^{-12} C/N)	593	374	149	-33
d_{31} (10^{-12} C/N)	-274	-171	78	23
k_{33}	0.75	0.71	0.48	0.15
k_{31}	0.39	0.31	0.21	0.12
Curie Temp ($^{\circ}$ C)	195	365	120	110

TABLE 1. Coefficients of common piezoelectric materials [18, 11]

typically takes the form of a coil and the electricity is generated by either the relative movement of the magnet and coil, or because of changes in the magnetic field. In the former case, the amount of electricity generated depends upon the strength of the magnetic field, the velocity of the relative motion and the number of turns of the coil.

The damping coefficient arising from electromagnetic transduction c_e can be estimated from equation (10) [8].

$$c_e = \frac{(n\ell B)^2}{R_{load} + R_{coil} + j\omega L_{coil}} \quad (10)$$

where N is the number of turns in the generator coil, ℓ is the side length of the coil (assumed square), and B is the flux density to which it is subjected and R_{load} , R_{coil} , and L_{coil} are the load resistance, coil resistance and coil inductance respectively. Equation (10) shows that R_{load} can be used to adjust c_e to match c_p and therefore maximise power, although this must be done with the coil parameters in mind. It can be shown that the maximum average power delivered to the load can be found from equation (11) [25].

$$P_{load\ max} = \frac{mA^2}{16\zeta\omega_n} \left(1 - \frac{R_{coil}}{R_{load}}\right) \quad (11)$$

Electrostatic Transduction Electrostatic transduction exploits the relative movement between two dielectrically isolated electrodes (capacitor). These plates are charged by periodic connection to a voltage source or by the use of electrets. The energy stored in a capacitor, with plate charge Q and potential difference V , is given by:

$$E = 0.5QV = 0.5CV^2 = 0.5Q^2/C \quad (12)$$

If the charge on the plates is held constant the perpendicular force between the plates is given by:

$$F = 0.5Q^2d/\epsilon A \quad (13)$$

If the voltage between the plates is held constant the perpendicular force between the plates is given by:

$$F = 0.5\epsilon AV^2/d^2 \quad (14)$$

The work done against the electrostatic force between the plates provides the harvested energy.

In MEMS the separation between the two plates is typically very small (nm to μm range). Electrostatic generators can be classified into three types [22]:

1. In plane overlap varying (Fig. 2)
2. In plane gap closing (Fig. 3)
3. Out of plane gap closing (Fig. 4)

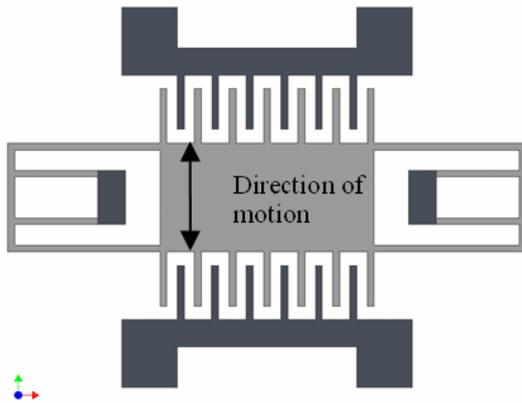


FIGURE 2. *In plane overlap varying*

Note both in plane configurations create two variable capacitors with the capacitances 180° out of phase. The three approaches can be operated either in charge constrained or voltage constrained cycles. In general the voltage constrained offers more energy than the charge constrained approach. However, by incorporating a capacitor in parallel with the energy harvesting capacitor, the energy from the charge constrained system can approach that of the voltage constrained system as the parallel capacitance approaches infinity. This parallel capacitor effectively constrains the voltage on the energy harvesting capacitor [19]. Maximum power generation occurs for very small dielectric gaps in which case the source impedance can be very high resulting in poor power delivery.

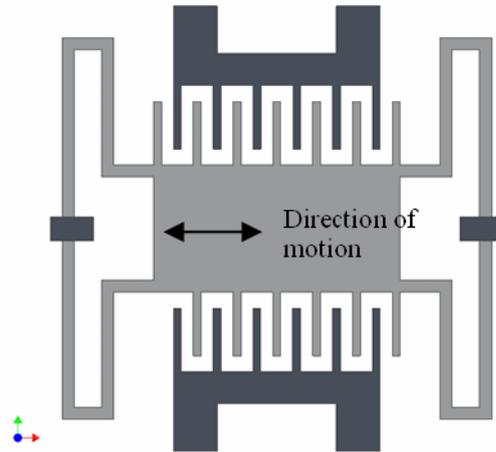


FIGURE 3. *In plane gap closing*

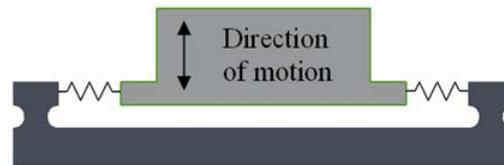


FIGURE 4. *Out of plane gap closing*

Table 2 provides the electrostatic force variation for the three configurations where x is the displacement of the inertial mass [6]. For a high damping configuration the electrostatic damping force has to be counter balanced almost entirely by the mechanical spring force.

Structure	Charge constrained	Voltage constrained
In plane overlap varying	$F_e \frac{1}{x^2}$	F_e constant
In plane gap closing	$F_e x$	$F_e \frac{1}{x^2}$
Out of plane gap closing	F_e constant	$F_e \frac{1}{x}$

TABLE 2. *Electrostatic force variation for the three configurations*

4 State of the art

The state of the art in electromagnetic and piezoelectric vibration power generators has been achieved in the VIBES project.

Electromagnetic Generator The micro electromagnetic generator was designed to operate at 50 Hz targeting electrical equipment driven by mains electricity [3]. The generator design is based upon a cantilever beam spring element and uses miniature discrete components fabricated using a variety of conventional manufacturing processes. This enables the generator to exploit the advantages of bulk magnetic material properties and large coil winding density thereby demonstrating very useable levels of power from a compact design. The design uses four high energy density sintered rare earth Neodymium Iron Boron (NdFeB) magnets manually bonded to the top and bottom surfaces of a cantilever beam. The magnetic poles are aligned as shown in Figure 5. The magnetic circuit is completed by zinc coated mild steel keepers which couple the flux between top and bottom magnets. This arrangement produces a concentrated flux gradient through the stationary coil as the magnets vibrate.

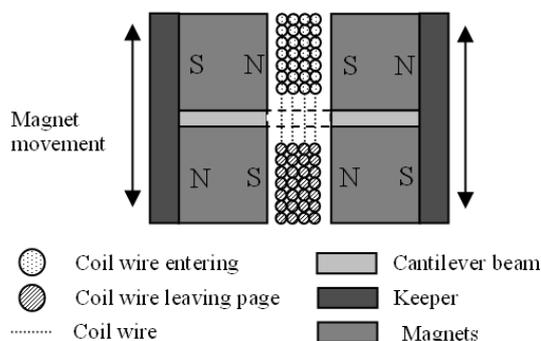


FIGURE 5. Cross section through the four magnet arrangement

Additional mass is added to the generator in the form of wire eroded tungsten alloy blocks attached to the free end of the cantilever beam. The tungsten alloy has a density of 18.1 g/cm^3 providing a compact inertial mass. Adjustment of the generator's resonant frequency can be achieved by altering the cantilever beam length. ANSYS modal analysis predicted an adjustment of 1.3 mm would produce a frequency range of 44-60 Hz. Therefore, a slot was etched into the beam at the clamping point to allow for manual adjustment of the beam length. Additionally, the coil has outer and inner diameters of 2.8 mm and 0.5 mm respectively. It is wound with $12 \text{ }\mu\text{m}$ diameter copper wire achieving 2800 turns and a coil resistance of 2.3 k Ω . The cantilever beam assembly

was clamped onto the base using an M1 sized nut and bolt and a square washer. The square washer gives a straight clamped edge perpendicular to beam length. The coil was manually bonded to a semi-circular recess machined in the base. A drawing of the assembled generator is shown in Figure 6. With the aid of alignment jigs, a tolerance of better than 0.1 mm can be achieved with the manual assembly of the components. The practical volume, i.e. including the swept volume of the beam, is approximately 0.15 cm^3 , but this increases to 0.8 cm^3 when packaged with a protective lid and weighs 1.6 g.

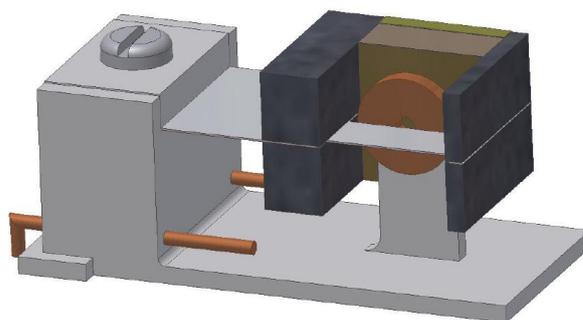


FIGURE 6. Micro cantilever generator

The generator performance was evaluated using an accelerometer feedback controlled shaker unit with a programmable resistive load. The tests involve sweeping through a range of frequencies with varying load resistances from 100 to 9 M Ω . The generator was subjected to a vibration acceleration of 0.6 ms^{-2} throughout the tests. The optimum power output of $58 \text{ }\mu\text{W}$ and a voltage of 1.12 V was achieved with a 15 k Ω load resistance. Connection to a voltage multiplier circuit and supercapacitor for energy storage results in a reduction in peak power to $37 \text{ }\mu\text{W}$ but the useable bandwidth increases to over 1 Hz (see Figure 7). This is due to the increased damping arising from the improved coupling provided by a capacitive load [26].

This increased damping means the generator can be driven up to an acceleration level of 1.7 ms^{-2} before reaching z_{max} at which point the power output is 120 μW rms. This device converts 51% of the energy from the mechanical domain to the electric domain.

Piezoelectric Generator The piezoelectric generators developed in the VIBES projects are MEMS fabricated using silicon micromachining and thin film deposition techniques for the piezoelectric material. These genera-

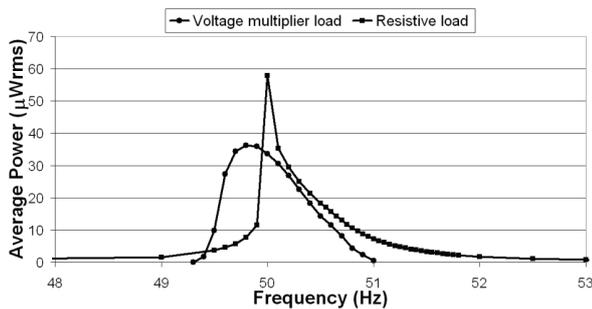


FIGURE 7. Maximum generated power into a resistive load of $15\text{ k}\Omega$ and into the voltage multiplier circuit, both at 0.6 ms^{-2} vibration

tors have been designed to match the higher frequency vibrations associated with a machine tool application ($>800\text{ Hz}$). The design is based upon a cantilever structure with a silicon inertial mass as shown in Figure 8.

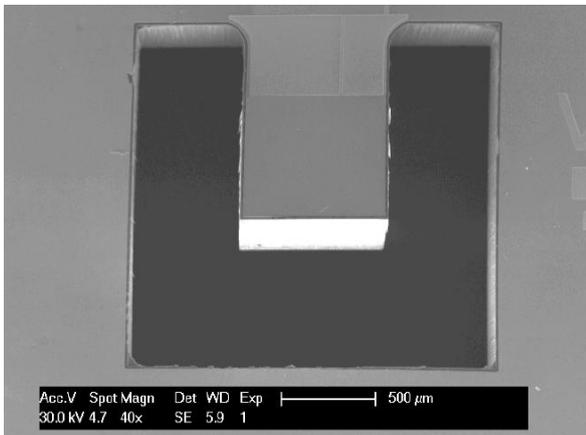


FIGURE 8. SEM photo of piezoelectric device

Two types of piezoelectric materials have been investigated: PZT and AlN. The AlN devices produced a maximum output power of $2\text{ }\mu\text{W}$ with voltage amplitude of about 1.6 V at 39 m/s^2 acceleration. The PZT devices have been fabricated with three different electrode arrangements and the maximum power of $1.4\text{ }\mu\text{W}$ at 19.6 m/s^2 acceleration amplitude was obtained with an IDT electrode arrangement.

Despite the low power output, the piezoelectric generator was combined with discrete electronic components and used to power a temperature sensing wireless system. The system was duty cycled to wake up from

sleep every 30 seconds, measure voltage on capacitor and if the voltage was higher than 2.1 V take a temperature sensor reading ($200\text{ }\mu\text{s}$) and transmit 5 bytes. This system was able to take measurements every 2 minutes.

5 Space applications

To the authors' knowledge, vibration energy harvesting has yet to be employed in space applications. Its potential is dependant upon the availability and nature of kinetic energy within potential application environments. Another consideration is the size of the generators. It is desirable for devices deployed in space to be as small and low mass as possible. However, in typical inertial vibration powered generators, the power output reduces linearly with mass, and therefore smaller devices produce less power output. Furthermore, the efficiency of microscale inertial generators is relatively poor at present. Electrostatic and piezoelectric devices are best suited to MEMS scale producing a few microwatts of power (depending upon excitation). Electromagnetics do not scale down well in size, which coupled with the reduced mass, leads to very low levels of power output (e.g. $<1\text{ }\mu\text{W}$). A beneficial aspect of operating in space will be the low pressure environment which removes gas damping effects.

One potential method for implementing this technology in space applications where no suitable vibrations are present is to provide an alternative method for actuating the generator. This has been demonstrated in the case of a piezoelectric cantilever beam which was coupled to radioactive source [16, 15]. The principle uses the radiated β particles to electrostatically charge a conductive plate on the underside of the cantilever. As the electrostatic field builds, the beam is attracted to the source until contact is made and the field dissipated. At this point, the beam is released to vibrate at its natural frequency and the kinetic energy harvested from the piezoelectric film. A schematic of the device is shown in Figure 9. Different material combinations, device geometries and radioisotopes can alter the output and characteristics of the generator. For example, a 1 cm square 0.5 millicurie thin film ^{63}Ni source with a half life of 100.2 years coupled to a 15 mm long, 2 mm wide silicon cantilever produced a peak power of $16\text{ }\mu\text{W}$ with a reciprocation period of 115 minutes [7]. This presents a novel and repeatable method for exciting the cantilevers vibrations and will work for very long time periods certainly more than alternative local power sup-

plies such as batteries. The power output is very periodic and, when averaged out over a given time period, very low (<1 nW).

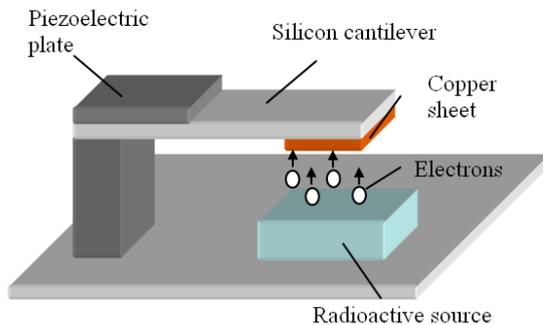


FIGURE 9. Radiation driven piezoelectric generator

6 Conclusions

This paper has reviewed vibration energy harvesting used to date to provide a localised power supply for wireless applications. The three transduction mechanisms employed to date to convert mechanical energy into electrical were presented along the characterising equations that highlight the important design parameters. Generator design and performance is inexorably linked to nature of the environmental excitation energy and successful implementation requires accurate knowledge of the application. Its suitability for space applications will be determined by the nature of the available kinetic energy. The radioactively coupled device may present an opportunity where suitable vibrations do not exist.

Acknowledgments

The authors would like to acknowledge funding for the VIBES project from the European Union, and the VIBES consortium, in particular our colleagues at the Tyndall National Research Centre, TIMA and EPFL.

References

- [1] R. Amirtharajah and A. P. Chandrakasan. Self-powered signal processing using vibration-based power generation. *IEEE J. of Solid State Circuits*, 33:687–695, 1998.
- [2] H. Baudry. Screen-printing piezoelectric devices. In *Proc. 6th European Microelectronics Conference*, pages 456–463, London, UK, 1987.
- [3] S. P. Beeby, R. N. Torah, M. J. Tudor, P. Glynn-Jones, T. O Donnell, C. R. Saha, and S. Roy. Micro electromagnetic generator for vibration energy harvesting. *Journal of Micromechanics and Microengineering*, 17:1257–1265, 2007.
- [4] S. P. Beeby, M. J. Tudor, and N. M. White. Energy harvesting vibration sources for microsystems applications. *J. Meas. Sci. Technol.*, 17:R175–R195, 2006.
- [5] K. Butt, A. Burstein, D. Chang, M. Dong, M. Fielding, J. Ho, J. Kaiser, E. Kruglick, F. Lin, T. H. Lin, H. Marcy, R. Mukai, P. Nelson, K. S. J. Pister, G. Pottie, H. Sanchez, O. M. Stafsudd, K. B. Tan, C. M. Ward, S. Xue, and J. Yao. A distributed, wireless mems technology for condition based maintenance. In *Proc. Integrated Monitoring, Diagnostics and Failure Prevention Conference*, pages 373–380, Mobile, Alabama, USA, 1996. Society of Machine Failure Protection Technology (MPFT).
- [6] G. Despesse, T. Jager, J. J. Chaillout, J. M. Léger, A. Vassilev, S. Basrou, and B. Charlot. Fabrication and Characterization of High Damping Electrostatic Micro Devices for Vibration Energy Scavenging. In *Proc. Design, Test, Integration and Packaging of MEMS and MOEMS*, pages 386–390, 2005.
- [7] R. Duggirala, H. Li, A. M. Pappu, Z. Fu, A. Aspel, and A. Lal. Radioisotope micropower generator for CMOS self-powered sensor microsystems. In *Proc. PowerMEMS*, pages 133–136, Kyoto, Japan, 2004.
- [8] M. El-hami, P. Glynn-Jones, N. White, M. Hill, S. Beeby, E. James, A. Brown, and J. Ross. Design and fabrication of a new vibration-based electromechanical power generator. *Sensors & Actuators: A. Physical*, 92(1-3):335–342, 2001.
- [9] S. Elliot, R. Saba, and O. N. Baumann. Vibration power harvesting from head motion. In *Proc. 7th European Conference of Structural Dynamics (Eurodyn)*, Southampton, UK, July 2008.

- [10] R. Frank. Move over, bluetooth; zigbee is here. *Design News*, 4, 2004.
- [11] J. L. Gonzalez, A. Rubio, and F. Moll. A prospect of the of piezoelectric effect to supply power to wearable electronic devices. In *Proc. 4th Int. Conf. on Materials Engineering for Resources*, pages 202–207, Akita, Japan, 2001.
- [12] G. Görgе, M. Kirstein, and R. Erbel. Microgenerators for energy autarkic pacemakers and defibrillators: fact or fiction. *Herz*, 26:64–68, 2001.
- [13] S. A. Jacobson and A. H. Epstein. An informal survey of power mems. In *Proc. The Int. Symp. on Micro-Mechanical Engineering (ISMME)*, page K18, Japan, 2003.
- [14] P. B. Koeneman, I. J. Busche-Vishniac, and K. L. Wood. Feasibility of micro power supplies for mems. *IEEE J. Microelectomechanical Systems*, 6:355–362, 1997.
- [15] A. Lal and J. Blanchard. The daintiest dynamics. *IEEE Spectrum*, 41:36–41, 2004.
- [16] H. Li, A. Lal, J. Blanchard, and D. Henderson. Self-reciprocating radioisotope-powered cantilever. *Journal of Applied Physics*, 92:1122–1127, 2002.
- [17] A. J. Lovinger. Ferroelectric polymers. *Science*, 220:1115–1121, 1983.
- [18] M. Matoroc. *Piezoelectric ceramics databook for designers*. Morgan Matoroc Ltd, 1996.
- [19] S. Meninger. *A Low Power Controller for a MEMS Based Energy Converter*. PhD thesis, Massachusetts Institute of Technology, 1999.
- [20] P. Mitcheson, B. Stark, P. Miao, E. Yeatman, A. Holmes, and T. Green. Analysis and optimisation of MEMS onchip power supply for self powering of slow moving sensors. In *Euroensors: 17th European conference on sensors and actuators*, pages 48–51, Guimaraes, Portugal, 2003.
- [21] V. Raghunathan, C. Schurgers, S. Park, and M. B. Skrivastava. Energy-aware wireless microsensor networks. *IEEE Signal Processing magazine*, 19:40–50, 2002.
- [22] S. Roundy, P. K. Wright, and K. S. J. Pister. Micro-Electrostatic Vibration-to-Electricity Converters. *Fuel Cells (methanol)*, 220:22, 2002.
- [23] S. Roundy, P. K. Wright, and J. Rabaye. A study of low level vibrations as a power source for wireless sensor nodes. *Computer Communications*, 26:1131–1144, 2003.
- [24] T. Starner and J. A. Paradiso. Human generated power for mobile electronics. In C. Piguet, editor, *Low Power Electronics Design*. CRC Press LLC, Boca Raton, Florida, 2004.
- [25] N. G. Stephen. On energy harvesting from ambient vibration. *Journal of Sound and Vibration*, 293:409–425, 2006.
- [26] R. N. Torah, P. Glynne-Jones, M. J. Tudor, and S. P. Beeby. Energy Aware Wireless Microsystem Powered By Vibration Energy Harvesting. In *Proc. of PowerMEMS*, pages 323–326, Freiburg, Germany, 2007.
- [27] T. von Büren, P. Lukowicz, and G. Tröster. Kinetic Energy Powered Computing: an Experimental Feasibility Study. In *Proc. 7th IEEE Int. Symposium on Wearable Computers (ISWC)*, pages 22–24, White Plains, NY, 2003.



Acta Futura 3 (2009) 63-69

DOI: 10.2420/AF03.2008.52

**Acta
Futura**

Biomass Based Fuel Cells - Application to Manned Space Exploration

A. HALME*

Dept. of Automation and Systems Technology, Helsinki University of Technology

Abstract. Long-term energy-demanding operations in remote off-the-grid locations, like in space exploration, require small, lightweight energy storage and power sources that are able to remain functional over long periods of time. In manned flights human secretions represent a potential source of methane and hence hydrogen, which can be used as fuel in fuel cells, or in propellant gas mixtures. Processing those secretions together with other biological waste provides fuel from available resources thus reducing fuel transportation from Earth and contributing to waste disposal in long missions, like Mars exploration. Fuel cells, in particular biocatalyzed fuel cells, offer a potential solution to the problem of generating electricity from the available energy with the aid of bacteria. They convert readily available substrates from renewable sources such as cereal materials, vegetable, fruits, fish meat and even human waste, to electricity and useful by-products such as water. Since the biocatalyzed fuel cells use concentrated sources of chemical energy, they can be small and lightweight, which is a crucial matter in the space application. There are in principle three ways to convert the biological energy available into electricity and heat. One way is to use traditional anaerobic digestion in order to produce methane and then further process it to electricity/heat by

the aid of high a temperature fuel cell like SOFC. Both of these technologies are already available at a practical level, although application to the space environment needs further development. The second way is to use a bacterial biological fuel cell for direct electricity production. This technology is fairly new, but has already been the subject of experiments in several practical tests. The third way is to use biocatalyzed electrolysis or reforming to produce hydrogen directly, which can then be further processed into electricity with the aid of a low temperature fuel cell, like PEM. This is a very new, but promising innovation, which still needs research to prove it's feasibility. This paper makes a short introduction to the underlying technology and compares the energy balance of the two first mentioned ways to recover electricity from biological waste recycling in an imaginary case, where six astronauts live in a micro ecological life supporting system during their Mars mission.

1 Introduction

The planetary objects to be explored in future have very different environment from our Earth. Some planets are covered with fluid; some moons or asteroids are icy or snowy; Some have (practically) no water and oxygen like our target planet - Mars. It means that most of the

*E-mail address: aarne.halme@tkk.fi

energy, food and water sources should be transported from Earth or carried with the spacecraft with the exception of the energy received from solar and wind energy, and some vegetables which could be grown in the spacecraft and on Mars. In order to reduce the burden on the spacecraft, it will be very important to have a micro ecological life support system especially in the landing mission phase on Mars. The detailed information can be found in the MELISSA project of ESA [3].

On the other hand, during a two and half-year trip to Mars, according to one estimate, a crew of six humans will generate more than six tons of solid organic waste – much of it secretions. So what do we do with all that? Right now, astronaut waste gets shipped back to Earth. But for longterm exploration, it is good to recycle it, because it holds resources that astronauts will need. It will provide pure drinking water. It will provide fertilizer. Or a part of the biomass can be converted to electricity to help with the energy balance of the mission facilities, as we will see later on.

Human secretions represent a potential source of methane and hence hydrogen, which can be used as fuel in a fuel cell, or in propellant gas mixtures. Such a process would at the same time provide fuel from available resources thus reducing fuel transportation from Earth and in addition contribute to waste disposal. The process of converting human excrement and vegetable residues into methane, carbon dioxide and other gases is anaerobic digestion. It is a well established process. It occurs naturally wherever high concentrations of wet organic matter accumulate in the absence of dissolved oxygen. The process takes place over a wide range of temperatures and with moisture content from 60 % to 90 % [6].

Fuel cell technology is becoming more and more important especially because of the high petroleum price nowadays. Fuel cells provide a range of critical benefits that no other single power technology can match [8].

A fuel cell uses the chemical energy of hydrogen and oxygen to directly produce water, electricity, and heat. They are therefore inherently clean and efficient and are uniquely able to address the issues of environmental degradation and energy security. They are also safe, quiet and very reliable. Fuelled with pure hydrogen, they produce zero emissions of carbon dioxide, oxides of nitrogen or any other pollutant. Even if fuelled with fossil fuels as a source of hydrogen, noxious emissions are orders of magnitude below those for conventional equipment. They offer significant improvements in energy efficiency as they remove the intermediate step of

combustion and mechanical devices such as turbines and pistons. Unlike conventional systems, they operate with high efficiency at part load. Also, unlike conventional plants, their high efficiency is not compromised by the small sizes. High efficiency saves fuel and reduces CO₂ emissions.

Fuel cells can use hydrogen derived from a variety of sources, including natural gas and coal, and renewable substances such as biomass or, through electrolysis, wind and solar energy. Fuel cells offer an opportunity to customers with a value-added energy service at a lower overall cost that is not subject to the same competitive or regulatory pressures as conventional electrical supply.

Biocatalyzed fuel cells (BFC) use biocatalysts instead of metallic catalysts used in chemical fuel cells. Biocatalysts could be micro-organism(s) or enzyme(s). The fuel is in general an organic substrate. In biocatalyzed fuel cell the biocatalysts participate in the electron transfer chain between the fuel substrates and the electrode surfaces. Both the fuel producing reaction and the electrode reaction take place in the same container. The biocatalyzed oxidation of organic substances by oxygen or other oxidizers at two-electrode interfaces provides means for the conversion of chemical energy to electricity. Abundant organic materials such as methanol, organic acids or glucose and even organic waste, like vegetable residues and human excrements, can be used as substrates for the oxidation process.

Both methane and hydrogen are potentially good fuels for the fuel cell systems. Biologically produced methane could also be chemically reformed with a miniature reformer to produce hydrogen for the fuel cell. Atmospheric oxygen or peroxide, H₂O₂, can act as the oxidant being reduced in the electricity producing process within the fuel cell. In space exploration the fuel transportation from Earth should be minimized. A good solution for the electricity production is through a biocatalyzed fuel cell using organic waste, food residues and human excrements, as fuel. Since the composition of the substrate is quite complex, an enzymatic fuel cell may not be the best choice. One type of enzyme can only use one type of substrate and each enzyme usually has a specific optimum condition. Thus, the better solution would be a microbial fuel cell system, which is a more robust choice in this case.

A review of biocatalyzed fuel cells giving a good overview of the recent research can be found in [1].

Another interesting recent discovery is biocatalyzed electrolysis to produce hydrogen from waste materials [7]. This a very new promising technology, which uses

the same principle as biocatalyzed fuel cells. The only difference is that instead of taking out electricity from the cell we connect a reverse voltage to its terminals, close the cathode for oxygen and allow the protons to be reduced at the cathode to produce hydrogen.

The case study reported in this paper is a shortened version of results from ARIADNA AO/1-4532/03/NL/MV [9] project, where we investigated two different routes for producing electricity by fuel cell systems from organic waste, in particular from human excrements, during an imaginary manned exploration of Mars. One of the routes is based on producing methane first with the aid of digestion and subsequently transforming it to electricity by a chemical fuel cell system. The second "direct" route is based on a biocatalyzed fuel cell system. A trade-off between the corresponding mass and energy balances is performed and discussed.

2 Fuel Cells - A Short Overview

There are various types of fuel cell systems available nowadays. Fuel cells are classified, according to the different features of the fuel cells, into: polymer electrolyte membrane fuel cells (PEM), alkaline fuel cells (AFC), phosphoric acid fuel cells (PAFC), molten carbonate fuel cells (MCFC), solid oxide fuel cells (SOFC), direct methanol fuel cell (DMFC) and biological fuel cells (BFC). They all operate with the same generic principle as shown in Figure 1.

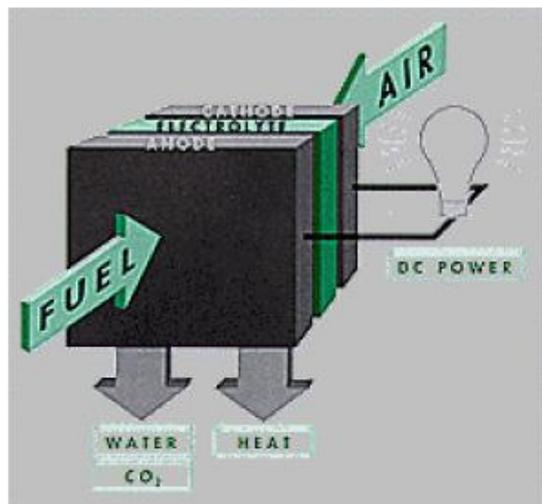


FIGURE 1. General principle of a fuel cell.

Chemical fuel cells can be further divided into categories according to their operating temperature [2].

Low temperature fuel cells:

- PEMFC
- DMFC
- AFC
- PAFC

High temperature fuel cells:

- MCFC
- SOFC

Most fuel cells operate with hydrogen gas. The exceptions are DMFC, which operate with liquid methanol, and high temperature cells, which also operate with more complex fuels, like natural gas/methane, CO, or even diesel. The reactions at the electrodes are different even if the overall reaction does not change (Figure 2). AFC is the oldest technology, used already in the 60's during the Apollo program.

Today development priorities are:

- PEM in car industry and small CHP-plants
- DMFC in electronics etc. applications
- high temperature fuel cells for CHP and larger power station applications

Fuel Cell	Anode Reaction	Cathode Reaction	Conducting Ion
Alkaline Fuel Cell (AFC)	$H_2 + 2(OH)^- \rightarrow 2 H_2O + 2e^-$	$1/2 O_2 + H_2O + 2e^- \rightarrow 2(OH)^-$	$(OH)^-$
Proton Exchange Membrane Fuel Cell (PEMFC)	$H_2 \rightarrow 2H^+ + 2e^-$	$1/2 O_2 + 2H^+ + 2e^- \rightarrow H_2O$	H^+
Direct Methanol Fuel Cell (DMFC)	$CH_3OH + H_2O \rightarrow CO_2 + 6H^+ + 6e^-$	$1/2 O_2 + 2H^+ + 2e^- \rightarrow H_2O$	H^+
Phosphoric Acid Fuel Cells	$H_2 \rightarrow 2H^+ + 2e^-$	$1/2 O_2 + 2H^+ + 2e^- \rightarrow H_2O$	H^+
Molten Carbonate Fuel Cells	$H_2 + CO_3^{2-} \rightarrow H_2O + CO_2 + 2e^-$	$1/2 O_2 + CO_2 + 2e^- \rightarrow CO_3^{2-}$	$(CO_3)^{2-}$
Solid Oxide Fuel Cells	$H_2 + O^{2-} \rightarrow H_2O + 2e^-$	$1/2 O_2 + 2e^- \rightarrow O^{2-}$	O^{2-}

FIGURE 2. Electrode reactions of chemical fuel cells.

In contrary to the chemical fuel cell, which is already a 100 year old innovation, the biocatalyzed fuel cell is quite a recent one. The early studies are from the beginning of 90's. The basic idea is similar to PEM, but reactions take place in the liquid phase and are catalyzed biologically either by living microbes or enzymes. The

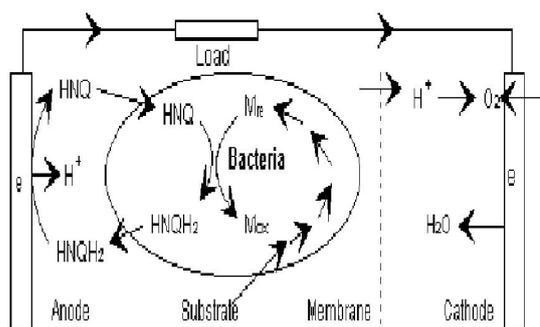


FIGURE 3. Principle of a bacteria fuel cell.

principle of a bacteria catalyzed fuel cell is shown in Figure 3.

The biological catalysts cannot normally make redox reactions directly with electrodes. This can be done only by the aid of specific mediator chemicals (HNQ in Figure 3). This together with the fact that biological substrates are usually very complex compounds usually makes the anode reactions more complex than in the case of chemical fuel cells. On the other hand, biocatalyzed fuel cells allow the use of a large variety of fuels - in principle all biologically decomposable substances can be used. All biological fuel cells operate in low temperatures close to ambient temperature. This is an advantage in many cases. The main disadvantage is a low power density of ~ 1 mW/cm² compared to those of ~ 60 mW/cm² in DMFC (operating also in liquid phase) and 300 - 400 mW/cm² of PEM and high temperature fuel cells operating in gas phase.

An interesting recent discovery closely related to biocatalyzed fuel cells is **biocatalyzed electrolysis**. The principle is shown in Figure 4. The same type of cell is used as in a PEM fuel cell, but the cathode compartment is closed and the external power source is added to aid proton reduction to hydrogen at cathode. The electrical power needed is less than that of the produced hydrogen, so it can be taken e.g from a fuel cell burning the hydrogen.

3 Application to a Manned Space Exploration Scenario

In the ARIADNA project [9] we studied two alternatives to recovery energy from recirculation of waste biomass during a manned Mars exploration lasting 2.5

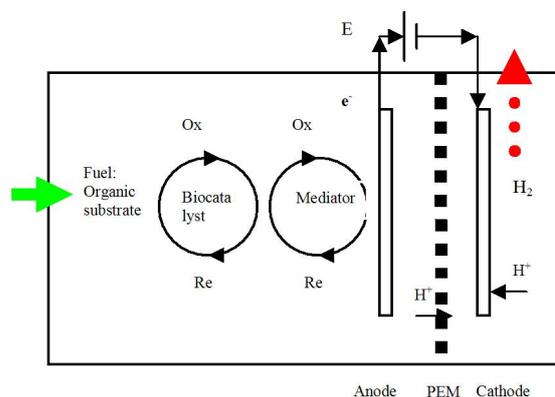


FIGURE 4. Principle of a biocatalytic electrolysis.

INPUT		OUTPUT	
O ₂ (kg)	0.83	CO ₂ (kg)	1.0
H ₂ O total ¹ (kg)	27.58	Faeces + toilet paper ² (kg)	
		Dry	0.053
		Wet	0.143
Food (grown) (kg)	1.0	Brine for urine (kg)	0.524
Food (packaged) (kg)	0.565	Brine for shower/handwash/sweat (kg)	0.254
		Plant biomass (from harvesting, cooking and left-overs) ³ (kg)	4.025
		Wet trash (paper, wipes, 10% humidity) (kg)	0.26
		Dry trash (tapes, filters, packaging, misc.) (kg)	0.6

¹ 97% of water is circulated. The remaining 3% goes along with brines.

² (0.03kg dry faeces only).

³ Wet. Includes 10% of left-overs and 30% processing waste.

TABLE I. Input and output of an astronaut per day, Extended Base, All plants menu, limited to items applicable for fuel-cell study.

years. Two alternatives were studied and compared technically including the calculated energy balances. One alternative was to start with the anaerobic, process and the biogas and finally produce electricity by the aid of chemical fuel cells. The other alternative was to use biological fuel cell technology, which provides a “direct way” to convert biomass to electricity. Both ways need a deeper analysis of the content of the biomass in the circulation. This in turn depends a lot on mission type, especially on the type of astronaut’s diet. In particular the amount of grown food is important, as inedible plant biomass plays a significant role in waste material circulation.

During transit to Mars the possibilities to grow food on the pace ship are very much different from those on the Mars base. Also in the Mars base there are several different ways to build the astronaut’s diet, which in turn has a great effect on amount and nature of circulating matter. The references [4] and [5] present 6 different diets of which one was selected to be a starting point for

INPUT		OUTPUT	
CO ₂ (kg)	0.735	O ₂ (kg)	0.534
H ₂ O (kg)	6.5	Edible food (kg)	1.0
Energy(light) ¹ (kW)	69.7	Non-edible biomass (kg)	4.0
Needed area (m ²)	26.8		
¹ Energy (light) 2.6 kW/m ²			

TABLE 2. *Input and output of plant field per day(per person), limited to items applicable for fuel-cell study, Extended Base, All Plants Menu.*

this study. The selected scenario is the “Extended Base, All Plants Menu”. When moving on to the “Transit to Mars”-menu the ratio of grown food decreases and the amount of packaged food increases. This increases the amount of packaging waste and decreases the amount of plant biomass waste. Two material flow models were characterized: input and output of a crew member, and input and output of plant growing facility. The third flow including input and output of the gasification/fuel cell systems can be characterised based on this information. For interest, the amounts of estimated inputs and outputs of an astronaut and the plant field per day are illustrated in Tables 1 and 2. In Table 3 we present a summary of the waste biomass components for a crew of a single and six astronauts.

	One person	6 people
Faeces		
Rate (kg wet/day)	0.150	0.9
Ash (kg/day)	0.0075	0.045
Biodegradable waste (kg dry/day)	0.03	0.18
Energy density (MJ/kg dry biodegradable waste)	11.8	
Energy (MJ/day)	0.354	2.124
Vegetable residues and others		
Rate (kg wet/day)	4.00	24.0
Biodegradable solid waste (kg dry/day)	1.22	7.32
Energy density (MJ/kg dry biodegradable waste)	17.5	
Energy (MJ/day)	21.35	128.1
Overall mass (kg wet/day)	4.150	24.90
Overall energy (MJ/day)	21.7	21.7
Overall solid biodegradable waste (kg/day)	1.25	7.50
Volume density (kg/m ³)	300.0	
Overall volume (liter)	4.17	25.0

TABLE 3. *A summary of the components of the waste biomass and its energy content.*

The overall energy balance of the Mars base is illustrated in Figure 5. The biomass recycling and energy recovering are indicated by green arrows. Figure 6 de-

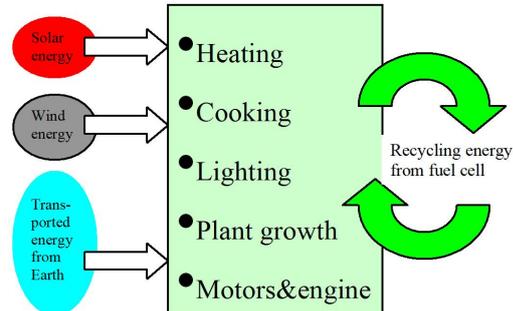


FIGURE 5. *The overall energy system.*

scribes the principle of the anaerobic digestion/chemical fuel cell process line and Figure 7 describes the biocatalyzed fuel cell process line. The efficiencies of individual components are rather well - known today. In spite of that, calculation of the energy balances are rather lengthy process. Details can be found in [9]. Estimations for the equipment volumes as well a preliminary design for the biocatalyzed fuel cell reactors can be also found there.

Due to the low power density of the reactor the required volume may be a problem, but otherwise the biocatalyzed fuel cell line seems to be a slightly better choice when comparing the energy balances and technical complexities of the lines.

The calculated energy balances turn out to be slightly positive in both cases as shown in Table 3. Also the mass balances are interesting to consider what happens to the life supporting system. All other substances except CO₂ can be directly reused. At least part of the carbon dioxide can be reused in the greenhouse, but we did not estimate how much.

4 Summary and conclusions

Biomass energy can be converted into electrical energy when recycling biomass waste in micro ecological life supporting system during long space flights. The net balance of recovery is not much but it is positive and seems to be slightly better when using biocatalyzed fuel cell technology than when using classical anaerobic digestion (reforming of biogas and chemical fuel cells). A new biocatalyzed electrolysis to produce hydrogen directly from biomass seems very promising and may bring a new dimension to this analysis. Chemical fuel

Anaerobic digestion/reforming/fuel cell	Biocatalyzed fuel cell
Mass balance	
Input: 7.5 kg biodegradable waste 6 kg oxygen Output: 1.5 kg methane (+4.1 kg CO ₂ + 1.9 kg compost) from AD process 4.1 kg Water + 3.3 kg CO ₂ from FC	Input: 7.5 kg biodegradable waste 3.5 kg oxygen (100% converted) Output: 2 - 3 kg compost 3 - 3.5 kg Water 5 - 5.5 kg CO ₂ from FC
Energy balance	
Input: 130 MJ/day in biodegradable waste Output: 26.2 MJ/day after AD process (20% efficiency) 5.2 - 7.8 MJ/ after FC system (20-30% efficiency) Overall energy efficiency: 4 - 6 %	Input: 130 MJ/day in biodegradable waste Output: 39 MJ/day from the BFC system 30 MJ/day consumed for the process 9 MJ/day (104 W) Overall energy efficiency: 6.9 %

TABLE 4. Calculated energy balances.

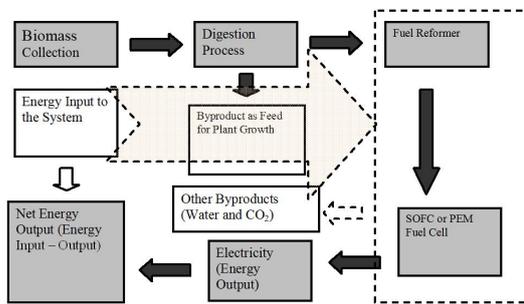


FIGURE 6. Energy recovery process when using methanization , gas reforming and SOFC or PEM fuel cells.

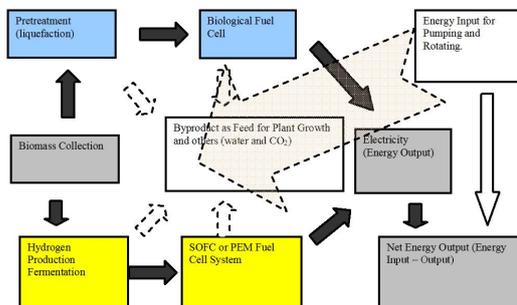


FIGURE 7. Energy recovery process when using methanization , gas reforming and SOFC or PEM fuel cells.

cell technology is already almost ready for applications, but biocatalyzed fuel cells are still in their early infancy phase. Because of their potential usefulness in long lasting manned space exploration missions their technology should be studied and developed more.

References

- [1] R. Bullen, T. Arnot, J. Lakeman, and F. Walsh. Biofuel cells and their development. *Biosensors and Bioelectronics*, 21(11):2015–2045, 2006.
- [2] EG&G Technical Services, editor. *Fuel Cell Handbook*. U.S. Department of Energy, Office of Fossil Energy, National Energy Technology Laboratory, 2008.
- [3] European Space Agency. Melissa project. <http://ecls.esa.int>, 2008.
- [4] NASA-JSC. Advanced life support systems modeling and analysis project. Technical report, Lockheed Martin Space Operations, Houston, Texas, 2001.
- [5] NASA-JSC. Advanced life support baseline values and assumptions. Technical report, Crew and Thermal Systems Division, Lyndon B. Johnson Space Center, Houston, Texas, 2004.
- [6] T. Pipoli. Feasibility of biomass based fuels cells for manned space exploration. In *Proc. Seventh European Space Power Conference*, Stresa, Italy, 2005.

- [7] R. Rozendal, H. Hamelers, G. Euverink, S. Metz, and C. Buisman. Principle and perspectives of hydrogen production through biocatalyzed electrolysis. *International Journal of Hydrogen Energy*, 31(12):1632–1640, 2006.
- [8] World Fuel Cell Council. Fuel cell europe. <http://www.fuelcelleurope.org>, 2008.
- [9] X. Zhang, T. Ylikorpi, G. Pepe, and A. Halme. Biomass based fuel cells for manned space exploration. Technical Report 04-2301, European Space Agency, the Advanced Concepts Team, 2005. Available online at www.esa.int/act.



Orbital energy of natural satellites converted into permanent power for spacecraft

J. PELÁEZ*

*Prof. of Aerospace Engineering, Technical University of Madrid (UPM), ETSI Aeronáuticos,**Pz. Cardenal Cisneros 3, E-28040, Madrid, Spain*

1 Introduction

Some of the most startling discoveries about our solar system have been made in the outer planets. Jupiter predominates among these planets for many reasons. Just like the Sun, it has its own Jovian system. The study of this system has advanced our understanding of the broader solar system for nearly four centuries. Five NASA spacecraft flew past Jupiter in the past decades: Pioneer 10 and 11, Voyager 1 and 2 and Galileo.

However, outer planet exploration has always been handicapped by a scarcity of power. Solar panels become rapidly ineffective further from the Sun. The solar intensity at Jupiter, 5 AU distant from the Sun, is only one twenty-fifth of its value at Earth. As a consequence, energy is a *scarce commodity* in this kind of missions and the total energy which will be consumed by the spacecraft should be transported onboard.

To tackle this problem *Radioisotope Thermoelectric Generators* (RTG's) have been used as the source power in missions to the outer planets. These devices use thermocouples to convert into electrical energy the heat released by the natural decay of a strongly radioactive element (usually Pu - 238). RTG's exhibit three char-

acteristics: *i)* the energy produced is expensive because the efficiency is low, ranging between 3% and 10% (at present, the estimated cost fluctuates from \$40,000 to \$400,000 per watt), *ii)* the potential risk is very high due to the management of very radiative substances and *iii)* the mass grows quickly with the energy produced. Table 1 summarizes power and mass of the RTG's used in several missions: the mass (in kg) is ~ 20% of the power (in watt).

Mission	Electrical W	Mass (kg)
Pionner 10	150	54.4
Voyager 1 y 2	470	117
Ulysses	290	55.5
Galileo	570	111
Cassini	800	168

TABLE 1. RTG's in past missions.

These are reasons why NASA's Jupiter Icy Moon Orbiter (JIMO) mission investigated the use of a nuclear-powered craft. The spacecraft was to be propelled by a new kind of ion thruster (HiPEP), and powered by a small fission reactor. Such a mission to the Jovian

*Corresponding author. E-mail address: j.pelaez@upm.es

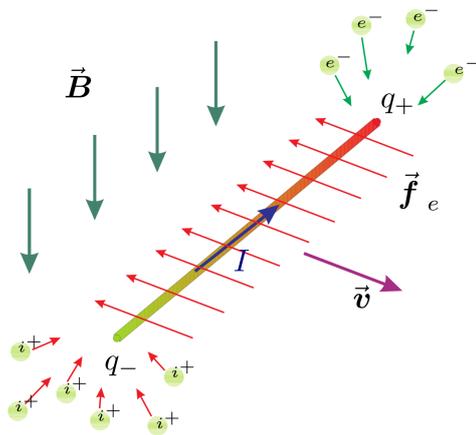


FIGURE 1. Conductive rod in a plasma environment moving in a magnetic field.

Galilean moons, initially planned for 2015, was finally canceled in 2005.

2 Electrodynamic tethers

A conductive rod moving in a magnetic field \vec{B} experiences a motional electric field given by

$$\vec{E} = \vec{v} \times \vec{B}$$

In the *vacuum*, a redistribution of surface charges takes place, leading to a vanishing electric field inside the rod. Thus, a steady state is reached with no motion of charged particles.

If the rod is moving inside a *plasma environment*, this picture changes drastically. The plasma electrons are attracted by the anodic end of the rod and the ions by the cathodic end. Some of these charges are trapped by the rod and produce a current I which flows inside the conductive material. The amount of current can be increased with the help of plasma contactors placed at the rod ends. Such a rod is, basically, an electrodynamic tether (ET).

The interaction between the tether current I and the magnetic field \vec{B} gives place to forces acting on the rod. They will brake (or accelerate) its motion. Their resultant, $\vec{f}_e = L(\vec{I} \times \vec{B})$ is the *electrodynamic drag (or thrust)*. There are two basic regimes for an ET: the generator regime (drag) and the thruster regime (thrust).

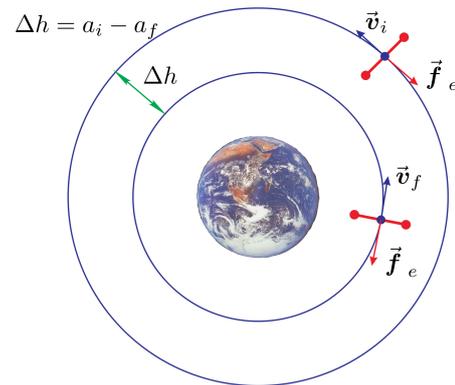


FIGURE 2. Deorbiting a satellite.

Let m be the mass of a satellite orbiting around a planet. When it descends from an initial orbit of radius a_i to a final orbit of radius a_f the mechanical energy lost in the process is given by

$$\Delta E = \frac{m\mu}{2a_f} \frac{\Delta h}{a_i}$$

where μ is the gravitational constant of the planet. This approximated result assumes the orbital system behaves as a point mass.

ET's can deorbit satellites in different scenarios (see [14, 2, 4, 5]) due to their capacity to produce *electrodynamic drag* \vec{f}_e when they operate in the generator regime. Such a *drag* brakes the spacecraft and provides the decay of the orbit. However, there are two essential requirements for the operation of an ET: 1) the magnetic field \vec{B} and 2) the plasma environment. Both are absent at the Moon where an ET can not be operated. However, the Earth and other planets, Jupiter for example, are appropriate for the use of ET's.

3 Power generation

When using an ET the mechanical energy lost in the descent process is dissipated through different mechanisms (see [8]). Roughly speaking, one part is spent bringing electrons from infinity to the tether. More dissipation takes place at the cathodic contactor of the tether and through the ohmic losses along the wire. Finally, a part is dissipated in *any interposed load* Z_T placed at the cathodic end of the tether.

This last contribution is, in fact, **useful energy, i.e., energy that can be used onboard to perform some task**

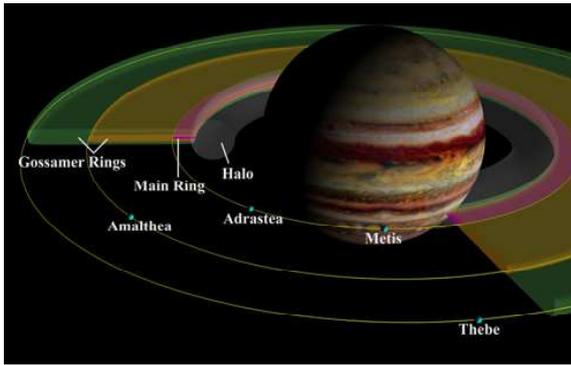


FIGURE 3. Jovian moonlets (from [1]).

(charging batteries, moving electrical engines or feeding some electronic circuit). Depending on the tether configuration and the environment conditions this contribution could reach 40% of ΔE ; values around 20% would be considered quite reasonable.

This property allows one to consider electrodynamic tethers as power sources, able to supply the whole system. In the generator regime they work converting mechanical energy into electrical energy and producing the required level of onboard energy; the bare tethers are particularly effective [13, 12]. As we show later on, in some missions a bare ET would provide much more energy than RTG's, enabling the utilization of more powerful instrumentation.

The essential idea of this work is to use an ET to obtain interesting amounts of power by deorbiting one of the Jupiter moonlets. Assume, for a moment, that the tether is joined to the moonlet with a cable; when the ET is switched on, *the electrodynamic drag will be transmitted to the moonlet through the tension of the cable and the deorbiting process will start.* The moonlet will be deorbited and therefore, its orbital radius will decrease as time goes on. This way the tether is able to recover a fraction of the mechanical energy lost by the moonlet during the deorbiting process.

Two aspects should be underlined. First of all, the masses of the moonlets are larger than 10^{16} kg. For example, the mass of Amalthea is $m = 2.09 \times 10^{18}$ kg and its orbital radius is $a_i = 181,300$ km. If we deorbit Amalthea only 1 mm ($\Delta h = 1$ mm) a large amount of energy is obtained

$$\Delta E \approx 4 \times 10^{15} \text{ J} \approx 1.1 \times 10^9 \text{ Kwh}$$

Thus, from a practical point of view, *the available energy*

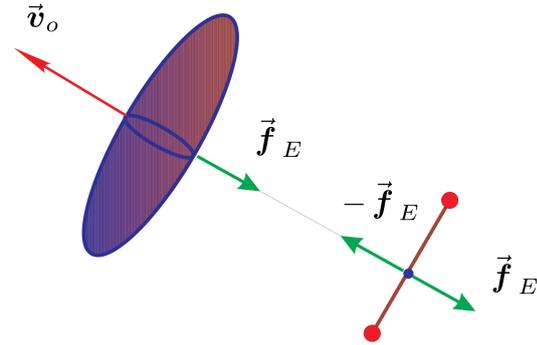


FIGURE 4. Deorbiting a Jovian moonlet.

is unbounded.

Second, the cable joining the S/C and the moonlet is not necessary; in effect, we can remove the cable if we place the S/C close enough to the moonlet. In such a case, the gravitational force of the moonlet substitutes for the cable tension and the deorbiting process takes place in the same way.

If we place the probe in an equilibrium position relative to the system Jupiter-moonlet the ET will carry on deorbiting the moonlet indefinitely. Adjusting the tether length, diameter and material it is possible to obtain electrical power in a sustained way at Jupiter's neighborhood.

An ambitious attempt like this has, obviously, obstacles. Without a doubt, one of the greatest challenges is the extreme Jupiter radiation environment. This is a serious constraint that can be considered as a significant challenge for current or near-term developing technologies. Another problem is the uncertainty associated with the *electronic plasma density* in the neighborhoods of Jupiter; this parameter is important in the operation of any ET because the tether current strongly depends on it. In the following we will not focus on these challenges, however.

4 Dynamics

A detailed dynamical analysis can be found in [11, 10]. The main aspects will be described here from a global point of view.

There are three bodies involved in the problem: Jupiter, the inner moonlet and the S/C. Since all the moonlets have circular orbits around Jupiter with very small inclination it is appropriate to consider the prob-

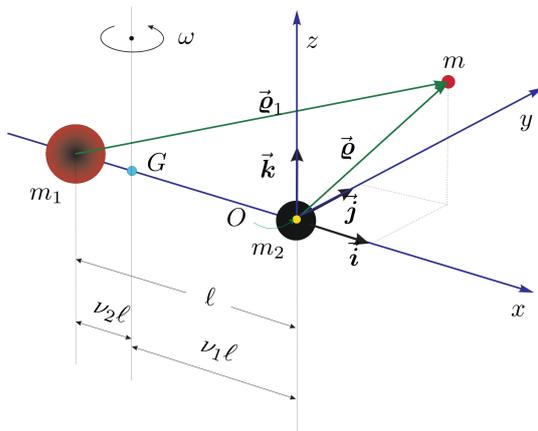


FIGURE 5. Reference frame.

lem as a generalization of the classical circular restricted three body problem (CRTBP), in a first approximation.

However, two significative differences should be emphasized: 1) electrodynamic forces are acting on the system and, 2) the attitude of the S/C should be considered since it plays an important role. Fig. 5 shows the *synodic frame* $Oxyz$ with origin at the center of mass of the Jovian moonlet, and where both primaries are at rest. The S/C must be placed in an equilibrium position relative to this frame.

Four main forces act on the system: the Jupiter gravity gradient, the Coriolis force, the gravitational attraction of the moonlet and the electrodynamic drag. Since the reduced mass ν of the moonlet is small, the Hill approach provides an appropriated model to tackle the dynamic problem.

We will take a *bare* electrodynamic tether (BET) because is more effective collecting electrons. The electrodynamic forces play, obviously, an important role. They depend on the external fields (magnetic field and plasma environment) and on the tether design (material, cross section, length). Two non-dimensional parameters, (χ, ε) measure their influence:

$$\chi = \frac{I_m B_0 L}{m \ell \omega^2 \nu^{1/3}}, \quad \varepsilon = \frac{J_1 B_0}{I_s \omega^2} \quad (1)$$

ELECTRODYNAMIC DRAG: the parameter χ is a measure of the electrodynamic drag. I_m is the averaged tether current, B_0 the magnetic field, L the tether length, m the mass of the S/C, ℓ the distance between both primaries, ω the angular rate of the synodic frame

and ν the reduced mass of the Jovian moonlet.

LORENTZ TORQUE: the parameter ε is a measure of the Lorentz torque at the center of mass of the S/C. I_s is the moment of inertia of the S/C relative to a straight line normal to the tether by the center of mass, and J_1 is the following integral

$$\frac{J_1}{\sigma E_m A_t L_*^2} = \int_0^{\ell_t} (\ell_t \cos^2 \phi - \sigma) i_e(\sigma) d\sigma \quad (2)$$

where σ is the non-dimensional distance from the upper tether end, $i_e(\sigma)$ the current profile (non-dimensional), A_t the cross section of the tether, σ the electrical conductivity of the material, E_m the component along the tether of the induced electric field, ϕ the mass angle (see [7]), $\ell_t = L/L_*$ a non-dimensional tether length and L_* is a characteristic length typical from the bare ET's given by:

$$L_* = \frac{(m_e E_m)^{1/3}}{e 2^{7/3}} \left(3\pi \frac{\sigma h_t}{n_\infty} \right)^{2/3}$$

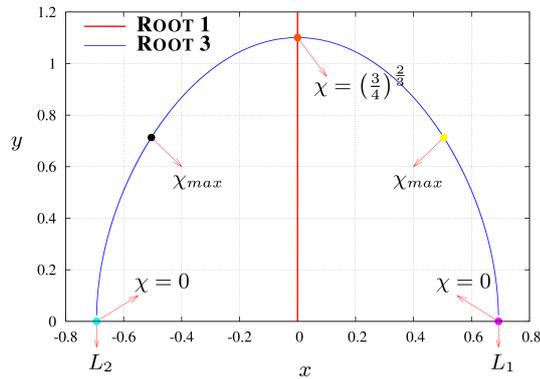
which essentially depends on the electronic plasma density n_∞ of the environment and the transversal distance h_t ; h_t is the radius for a round section, and the thickness for a tape (e electron charge, m_e electron mass).

A detailed analysis (see [11, 10]) provides two families of equilibrium positions of the S/C which depend on the value of χ . In the analysis the tether is self-balanced, that is, we select the mass angle ϕ in order to get a zero Lorentz torque (see [6, 9]). For both families the tether lies into the orbital plane of the moonlet. Fig. 6 shows the position of the center of mass for both families. We call *the main set* to the equilibrium positions close to the moonlet orbit (red line). A linear analysis shows stability when $\chi < 0.115$ for the main set; when $\chi > 0.115$ there are different instability regions and the tether can be operated in a reliable way by using a simple feedback control law (see [10]).

Fig. 7 shows the relative position of the tethered system in the synodic frame. In two moonlets (Amalthea and Thebe) the BET is just in front of the moonlet. In the others (Adrastea and Metis) the BET is just behind the moonlet. This is due to the different position of the moonlets relative to the stationary Jovian orbit.

5 External fields

MAGNETIC FIELD: in the neighborhoods of Jupiter the magnetic field is clearly dipolar and its polarity is just the


 FIGURE 6. Equilibrium positions (x, y) .

opposite to the polarity of the geomagnetic field. As a consequence, it can be modeled as

$$\vec{B}(\vec{r}) = \mu_m \frac{R_J^3}{r^3} (\vec{u}_m - 3(\vec{u}_m \cdot \vec{u}_r) \vec{u}_r) \quad (3)$$

where $\mu_m = 4.27 \cdot 10^{-4}$ Teslas is the intensity of the dipole, $R_J = 71492$ km is the equatorial radius of Jupiter and $-\vec{u}_m$ is a unit vector in the direction of the dipole. Since the tilt of the dipole is small, $\beta = 9.6^\circ$, we assume a non-tilted dipole. This way the magnetic field, considering the Hill approximation, takes the value

$$\vec{B} = B_0 \vec{k}, \quad B_0 = \mu_m \frac{R_J^3}{\ell^3}$$

INNER PLASMASPHERE: we follow the model of Divine & Garret (see [3]). When $R_J < r < 3.8R_J$ (the inner plasmasphere) the electronic plasma density, in m^{-3} , is given by

$$n_\infty = 4.65 \cdot 10^{-6} \exp \left\{ \frac{r_0}{r} - \left(\frac{r}{H_0} - 1 \right)^2 (\lambda - \lambda_c)^2 \right\}$$

The parameters involved are

$$r_0 = 7.68R_J, \quad H_0 = 1.0R_J, \quad \lambda_c = 0.123 \cos(l - l_0)$$

where l and λ are the longitude and latitude, in Jupiter System III (1965), respectively, and $l_0 = 21^\circ$.

This model provides a quasi-constant value for the electronic plasma density at the orbits of the Jovian moonlets ($\lambda = 0^\circ$). Fig. 8 shows the variations of n_∞ with l for the four satellites. The variation between the

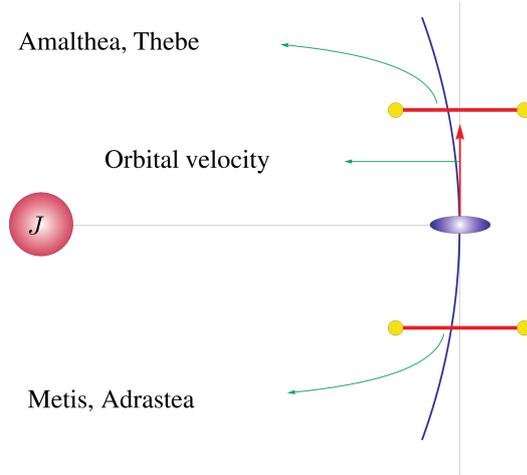


FIGURE 7. Relative position.

extreme values is lower than a 7% in all cases. This important fact permits to predict a small variation of the current collected in the BET without control.

We use another feature of the model of Divine & Garret: the most habitual ions in the region are: sulfur S^+ (about 70 %) and oxygen O^{++} (about 20 %).

6 Tether design

Fig. 9 shows an scheme of the BET. Let Z_T be the interposed load just placed at the cathodic end of the tether.

This load is essential and plays two complementary roles: 1) it is used to model the useful power obtained from the BET and 2) it permits the basic control of the system.

The useful power that can be obtained from the tether is given by

$$W_u = I_C^2 Z_T$$

However, the main parameter in the tether design is the ratio W_u/m_T where m_T is the

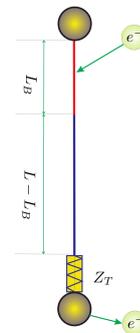


FIGURE 9. Tether scheme

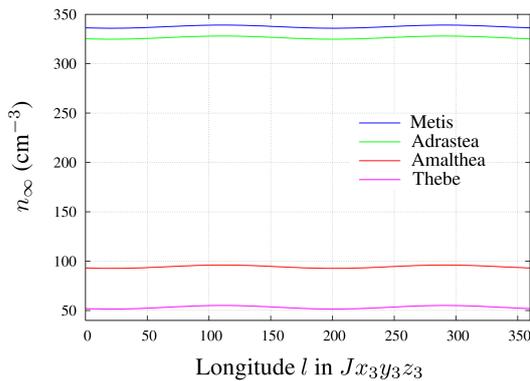


FIGURE 8. Electronic plasma density at the Jupiter moonlet orbits.

tether mass. It takes the value

$$\frac{W_u}{m_T} = \frac{\sigma E_m^2}{\rho_v} \cdot \Omega i_C^2(\ell_t, \Omega) \quad (4)$$

where ρ_v is the density of the material, $\Omega = Z_T/R_T$ is the non-dimensional form of the interposed load Z_T (here, $R_T = L/(\sigma A_t)$ is the electrical resistance of the tether, I_C the current at the cathodic end).

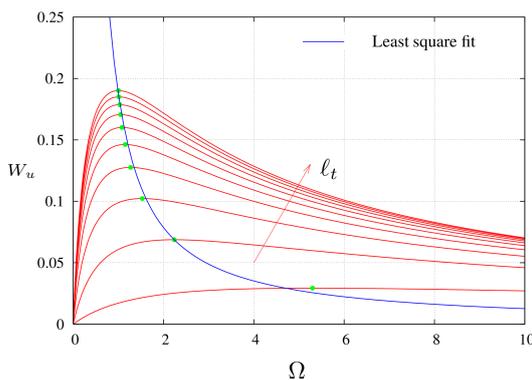


FIGURE 10. Useful power $W_u(\ell_t, \Omega)$; the picture shows the dependence on Ω for different values of $\ell_t = 0.5, 1.0, \dots, 5$.

In fact, assuming the BET is working in the OML regime, the electrodynamic performances of the tether are basically functions of only two parameters: the non-dimensional tether length ℓ_t and the non-dimensional interposed load Ω . Fig. 10 shows the dependence $W_u = W_u(\Omega)$ for different values of $\ell_t = 0.5, 1.0, \dots, 5$. It is clear that there is a line of max-

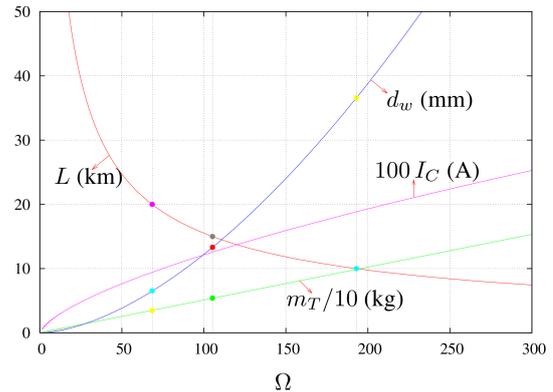


FIGURE 11. Optimum conditions at METIS ($h = 0.1$ mm, $W_u = 500$ w \sim Galileo S/C).

ima $\tilde{\ell}_t = \ell_t(\Omega)|_{\max}$. Along this line the tether performances are functions of only one parameter: the control parameter Ω .

TETHER MATERIAL: we select Aluminum since the ratio ρ_v/σ reaches a minimum value for it.

SELECTION OF THE JOVIAN MOONLET: for any BET, the current collection is governed by the external fields. Particularly, the electronic density of the surrounding plasma n_∞ and the magnetic field (through the component of the electric field in the direction of the tether E_m). High values of these parameters make easy to operate the system because benefit the electron collection by the tether. As a consequence we select **Metis as the Jovian moonlet where the tether will be operated.**

TETHER SECTION: first of all, we will select the shape of the *tether cross section*: a tape of thickness h and width d_w . This rectangular section is more appropriate than a circular one because the value of L_* is lower for the same cross section; a lower value of L_* provides a higher value of ℓ_t and, in general, the electron collection will be benefitted. As a consequence, **tape is better than wire.** We must select the value of h as small as possible. At present, it is possible to make tapes as thin as $h = 0.1$ mm, and this will be the thickness of our tape. Perhaps in the near future thinner tapes can be constructed.

USEFUL ENERGY: once the useful energy that the BET should provide is fixed, the tether length L , width d_w and mass m_T become functions of the control parameter Ω when we work on the line of maxima. Fig. 11 show this dependence for Metis and for a production of 500 watts of useful energy.

Figure 11 shows a wide range of **reasonable nominal**

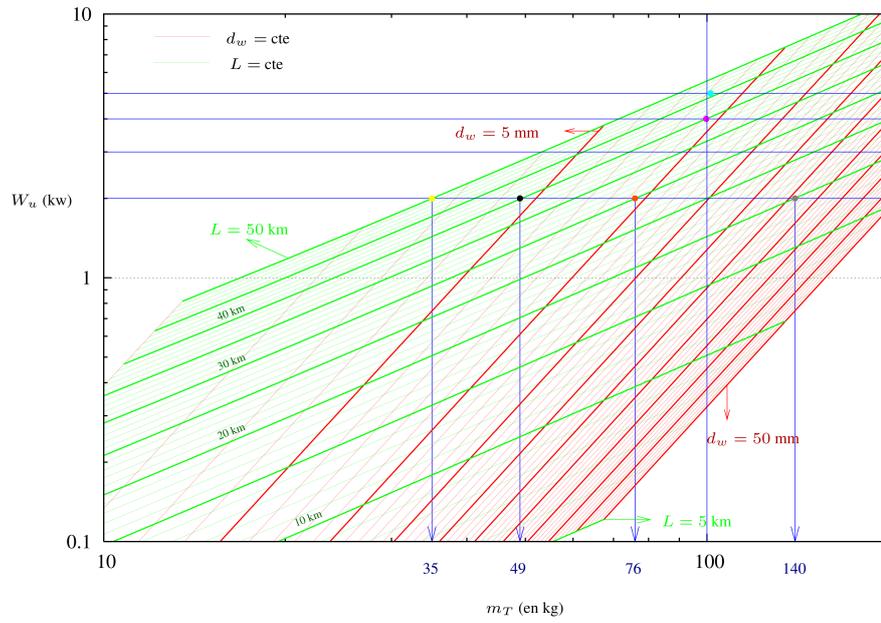


FIGURE 12. Different optimized configurations in METIS ($h = 0.1 \text{ mm}$).

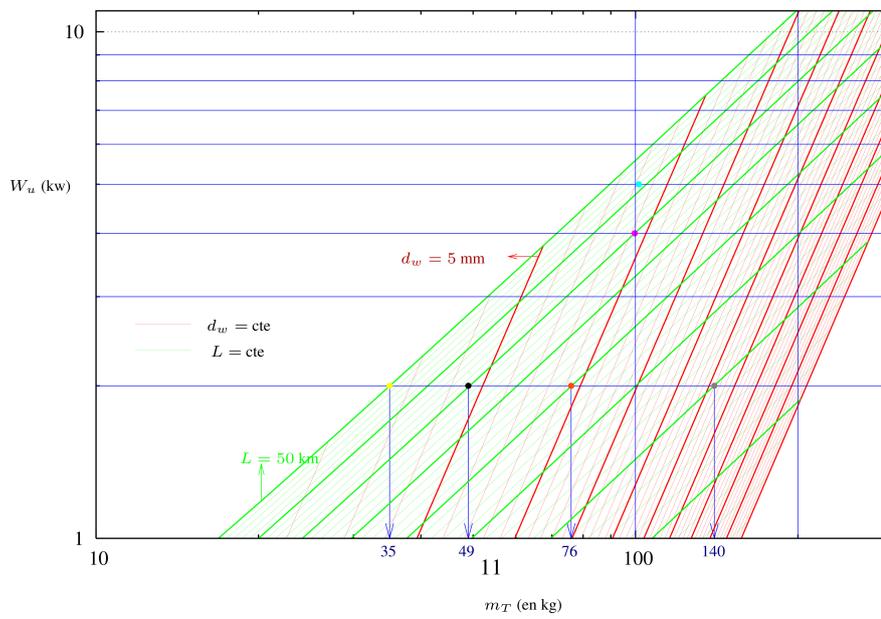


FIGURE 13. Zoom of the previous figure focused on the more interesting region.

values which provide the desired useful energy.

Obviously, the selection of one of these design should be made taking into account another factors not yet considered in the analysis. For example, the nominal value to be selected for Ω should involve a detailed study of the batteries charging process. These technological points should be clarified in the future, but for the moment they should be placed separately.

7 Optimization

To understand more deeply the optimization process we constructed figure 12 which correspond to a tape operated in Metis and about 0.1 mm thick. In the abscissa-axis figure shows the tether mass m_T (in kg); in the ordinate-axis figure shows the useful energy provided by the tether (in kilowatts); in both axes we use a logarithmic scale. On the figure two families of curves have been drawn: the green lines show the variation of the useful power W_u with the tether mass m_T when the tether length L is fixed; along these lines the only parameter which changes is the tether width d_w . The red lines show the variation of the useful power W_u with the tether mass m_T when the tether width d_w is fixed; along these lines the only parameter which changes is the tether length L . For the red lines, the slope is larger than for the green lines, this explains why the useful energy is more sensitive to the variations of the tether length L than to the variations of the tether width d_w . Obviously, changing the tether length and width simultaneously is possible to increase W_u keeping the tether mass m_T constant.

L (km)	d_w (mm)	m_T (kg)
50	2.7	35
40	4.7	49
30	9.5	76
20	26	140

TABLE 2. Some optimum configurations from figure 12 for $W_u = 2 \text{ kw}$.

Some important conclusions can be drawn from this figure. First of all, the BET provides more energy than the RTG's, for the same mass (around one order of magnitude more). For example, it is possible to obtain 500 w of useful power with a tether 17 km long and 10 mm width which weights 45.22 kg. The same level of energy

produced by RTG's require much more mass: 111 kg in the case of the Galileo spacecraft. But there is a range of values of the useful energy that cannot be reached with RTG's because of the prohibitive mass required and, however, it can be reached with an BET. For example, a tether 30 km long and 9.5 mm width which weights 76 kg is able to produce 2000 w (see Table 2). To obtain this amount of energy with RTG's the mass required would be, probably, about hundreds of kg. Finally, we underline that with tether masses about 200 kg is possible to reach useful powers about 10 kw.

8 Conclusions

From our analysis some conclusions can be drawn; we comment them briefly in what follows.

1) Significant amounts of energy can be obtained by deorbiting any of the inner Jovian moonlets (Metis, Adrastea, Amalthea or Thebe). From a practical point of view Metis, the innermost moonlet, is the preferable to be deorbitated with a bare, self-balanced, electrodynamic tether.

2) There exist equilibrium positions where the tether could be operated appropriately; some of them are stable and other unstable. At first sight, the operation of the probe in an stable equilibrium position would be preferable; however, the probe would be operated in an unstable equilibrium position with the help of a feedback control law.

3) The appropriate parameter to establish a control strategy is a *variable interposed load* placed at the cathodic end of the tether. In that strategy the interposed load plays two simultaneous roles: *i*) it simulates the electrical resistance associated with the batteries of the S/C and *ii*) it controls the tether current acting as a potentiometer (in series with the batteries).

4) For the same mass the BET is able to produce more energy than traditional RTG's. In fact, by increasing the tether mass it would be possible to obtain much more energy (one order of magnitud more). Thus, the bare tether would open new prospects unattainable with RTG's.

5) The onboard energy provided by the bare tether is cheaper than the energy provided by RTG's.

There are many more points involved in a mission as the one proposed in these pages. For example, how to face the strong radiation environment in the neighborhood of Metis, or how to place the probe precisely in the equilibrium position where it should be operated. Some

of these subjects will be studied in the near future.

References

- [1] F. Bagenal, T. Dowling, and W. McKinnon, editors. *Jupiter (The Planet, Satellites and Magnetosphere)*. Cambridge Planetary Science. Cambridge University Press, Cambridge, UK, 2004.
- [2] J. Corsi and L. Iess. Stability and control of electrodynamic tethers for de-orbiting applications. *Acta Astronautica*, 48:491–501, 2001.
- [3] N. Divine and H. B. Garrett. Charged particle distribution in jupiter’s magnetosphere. *Journal of Geophysical Research*, 88(A9):6889–6903, 1983.
- [4] L. Iess, C. Bruno, C. Ulivieri, U. Ponzi, M. Parisse, G. Laneve, G. Vannaroni, M. Dobrowolny, F. De Venuto, B. Bertotti, and L. Anselmo. Satellite de-orbiting by means of electrodynamic tethers - Part I: General concepts and requirements. *Acta Astronautica*, 50:399–406, 2002.
- [5] L. Iess, C. Bruno, C. Ulivieri, and G. Vannaroni. Satellite de-orbiting by means of electrodynamic tethers - Part II: System configuration and performance. *Acta Astronautica*, 50:407–416, 2002.
- [6] J. Peláez. Self-balanced electrodynamic tethers. In *Proc. of the AAS/ALAA Astrodynamics Specialist Conference*, Providence, Rhode Island, August 2004.
- [7] J. Peláez and M. Sanjurjo. Generator regime of self balanced electrodynamic bare tethers. *Journal of Spacecraft and Rockets*, 43(6):1359–1369, 2006.
- [8] J. Peláez and M. Sanjurjo. Power generation using self-balanced electrodynamic tethers in debris mitigation scenarios. In *Proc. of the AAS/ALAA Space Flight Mechanics Meeting*, Sedona, Arizona, January 2007.
- [9] J. Peláez, M. Sanjurjo, and J. Fontdecaba. Satellite deorbiting using a self balanced electrodynamic tether. In *Proc. of the 55th International Astronautical Congress*, Vancouver, Canada, October 2004.
- [10] J. Peláez and D. J. Scheeres. On the control of a permanent tethered observatory at Jupiter. In *Proc. of the AAS/ALAA Astrodynamics Specialist Conference*, Mackinac Island, Michigan, August 2007.
- [11] J. Peláez and D. J. Scheeres. A permanent tethered observatory at jupiter. dynamical analysis. In *Proc. of the AAS/ALAA Space Flight Mechanics Meeting*, Sedona, Arizona, January 2007.
- [12] J. R. Sanmartín, D. E. Hastings, and E. Ahedo. Systems analysis of electrodynamic tethers. *Journal of Spacecraft and Rockets*, 29:415–424, 1992.
- [13] J. R. Sanmartín, M. Martínez-Sánchez, and E. Ahedo. Bare wire anodes for electrodynamic tether. *Journal of Propulsion and Power*, 9(0):352, 1993.
- [14] G. Vannaroni, M. Dobrowolny, and F. De Venuto. Deorbiting of LEO satellites with electrodynamic tethers. *Proceedings Aerospace Sciences Meeting and Exhibit*, 2000.



Harnessing of the power of the solar wind particles captured in the Van Allen belts

E. K. KOLESNIKOV*, A. B. YAKOVLEV

Research Institute of Mathematics and Mechanics, St. Petersburg State University, Russia

Abstract. The feasibility of constructing a high-voltage electric generator (HEG) transforming kinetic energy of particles from the radiation belts into electric power is considered. The maximum specific power of the generator is theoretically evaluated for particular cases of setting it inside the natural radiation belts of the Earth (ERB) and in polar region. It is demonstrated that from the viewpoint of weight parameters, the suggested design of the HEG is quite competitive with power sources of low-thrust spacecraft operating on conventional principles.

1 Introduction

A power plant providing for power supply for the operation of onboard service devices as well as science and technology instrumentation is one of the main elements of any spacecraft. As of now, designing power plants transforming the natural energy of space medium into power supply is the most promising line of development of space power engineering. Power sources of this type include first of all various converters of electromagnetic radiation of the Sun (semiconductor photoelectric cells, thermoelectron, thermoionic, and thermoelectric converters). Along with manufacturing the power plants

that use the energy of electromagnetic radiation of the Sun, people in our country and abroad make research to develop concepts of radically new space electric power generators based on utilizing other types of energy available in space medium: energy of the Earth's magnetic field [13], energy of the solar wind plasma [3], and so on.

The radiation belts of the Earth and of other celestial bodies belong to the carriers of natural energy density comparable with the energy flux density of the solar electromagnetic radiation. In this paper, a principal feasibility of an electric generator converting kinetic energy of particles from the radiation belts and polar region into electric power is considered. The maximum specific power of the generator is theoretically evaluated for some particular cases. The fundamental part of this article results was published previously in [5].

2 The phenomenon of strong electrostatic charging of a shielded body in the flux of corpuscular radiation: principle of operation of a high-voltage power generator

It is well known that any body in open space is charged to a certain electric potential φ_p due to interactions with

*Corresponding author. E-mail address: kolesnikov_evg@mail.ru

cosmic plasma, with fluxes of high-energy charged particles, and also with photoemission fluxes and fluxes of secondary particles emitted by the surface of the body. In equilibrium, the value of this potential is determined by the condition of balance of charging currents. Under usual conditions the potential of a body in the space medium does not exceed a few volts. For example, the potential of a satellite in near-Earth space (NES) varies from negative values of order of some tenths of a volts (in low near-Earth orbits (LNO) with an altitude of several hundred of kilometers) up to positive values of about a few volts (in high near-Earth orbits (HNO) passing at altitudes of tens of thousands of kilometers). However, in some special cases a body can be charged up to considerably higher potentials. For example, when a satellite moving along HNO finds itself on a shadowed part of the orbit inside the rarefied high-temperature plasma of the plasma sheet, it can be charged to negative potentials of $\sim (1 - 10)$ kV [4]. Finally, strong electrostatic charging (SEC) of a body to negative potentials of order of $(10 - 1000)$ V is possible when this body is immersed into intense electron fluxes of natural origin (the fluxes of auroral electrons, electron fluxes of the radiation belt, the fluxes of photoelectrons from sunlit conjugate region of the ionosphere on a shadowed segment of the orbit). The smallness of the rate with which negative electric charges are emitted due to photoemission and balanced due to the flow of space plasma is the necessary condition of strong electrostatic charging of the body in the flux of high-energy electrons. Especially favorable conditions for SEC of a body by corpuscular fluxes are set up in the case when it is fully protected against solar shortwave radiation and space plasma by a special shield whose thickness is less than the mean free path of high-energy particles in its material. High-energy particles (in the case under consideration they can be represented by both, electrons and protons), decelerating inside a body, should charge it up to a certain potential U relative to the shield whose neutrality is maintained by charged particles from the ambient plasma. As will be shown below the maximum voltage U between the body and the shield can reach $\sim \bar{\epsilon}_r/e$ (where $\bar{\epsilon}_r$ is the mean energy of the particles of the corpuscular flux) and turns out to be very significant at large $\bar{\epsilon}_r$. The obvious application of the SEC phenomenon in a shielded body is the use of the electric potential U between the inner body and the shield in order to produce a electric current J_1 in the appliance load inserted between them. This current produces a power $P = J_1 U$. In this case the system will operate as a high-voltage electric generator

(HEG) converting the kinetic energy of the particles of a corpuscular flux into electric power.

As a first step in studying the efficiency of the above method of generating electric power in space, let us make a preliminary estimate of the possible HEG power for the case when the corpuscular fluxes producing the SEC phenomenon are represented by protons and electrons of the natural radiation belts of the Earth.

3 Specific power of a high-voltage electric generator in natural radiation belts of the Earth

The natural radiation belts of the Earth are extended regions of the near-Earth space that are characterized by intense fluxes of high-energy electrons and protons trapped by the Earth's magnetic field. The electron radiation belt consists of the inner and outer radiation belts, the gap between them is located near a magnetic shell with the McIlwain parameter $L = 3$. Unlike electrons, protons with energies ≤ 1 MeV, which give the main contribution to electric power produced by the HEG, occupy the entire region of trapping.

Let us first estimate possible values of the specific power of a HEG positioned in NES, where the flux of ERB protons is dominant. In accordance with the data of [2], this region is located on magnetic shells with values $2.5 \lesssim L \lesssim 4.5$.

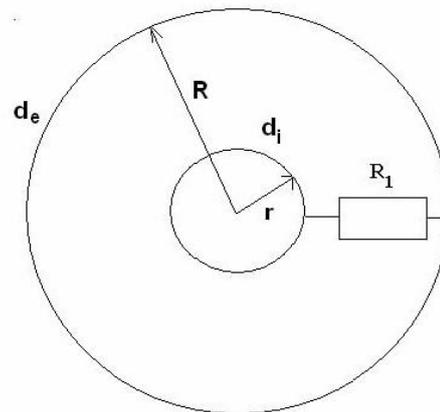


FIGURE 1.

For concreteness, let the generator be a system of two spherical concentric conducting shells. The internal shell has radius r and thickness d_i , while the external shell of radius R has thickness d_e , which is pre-

sumed to be substantially less than the mean free path of particles of the corpuscular flux in its material. The density of materials from which the internal and external shells are manufactured are denoted by ρ_i and ρ_e , respectively. The connection between the two spheres is supposed to have resistance R_1 (cf. Fig. 1). Some high-energy particles of the radiation belt, passing through the external shell and through the electric field inside the gap between the shells, are detained by the matter of the internal shell, thus imparting to it an electric potential relative to the external shell. The electric potential of the external shell is maintained close to zero due to the flux of charged particles with thermal energies from the ambient plasma. As a result, a high voltage $U \sim \bar{e}_r/e$ is established between the shells. This voltage produces an electric current in the load and thereby produces a corresponding power, i.e., the system operates as an electric generator.

Let us suppose that the high voltage U between the shells of the HEG is established as a result of partial absorption of the ERB proton flux by the material of the internal shell. In steady state, the voltage U is determined from the condition of balance of the currents charging the internal shell of the HEG,

$$J_{\text{abs}}^p(U) = J_1, \quad (1)$$

where $J_{\text{abs}}^p(U)$ is the total current of protons absorbed by the internal shell material, and the current flowing through the load, $J_1 = U/R_1$. In order to determine current $J_{\text{abs}}^p(U)$ let us first formulate the conditions under which a proton of energy E , which traverses the external shell of the HEG at an angle ϑ_1 compared to the generator's center direction, is absorbed by the matter of the internal shell of the HEG. If energy losses and scattering in the process of passing through the external shell can be neglected, the condition of reaching the internal shell may be written in the following form:

$$U_{\text{eff}} = \frac{M^2}{2m^2r} + U \leq E = \frac{mv^2}{2}, \quad (2)$$

where U_{eff} is the effective potential energy of the proton at the moment of entering the internal shell and $M = mvR \sin \vartheta$ is the angular momentum on the external boundary of the generator. It follows from (2) that the admissible angles of entering the generator are restricted by

$$\vartheta_1 \leq \vartheta_1^*(E) = \arccos \sqrt{1 - \frac{r^2}{R^2} \left(1 - \frac{eU}{E}\right)}. \quad (3)$$

If the direction of motion of a proton lies in the cone determined by condition (3) the proton should reach the internal shell. Having crossed the internal shell it will move at an angle ϑ_2 to the normal of the surface. This angle can be found from the integrals of energy,

$$\frac{mv^2}{2} \Big|_R = \frac{mv^2}{2} \Big|_r + eU, \quad (4)$$

and angular momentum

$$M \Big|_R = mvR \sin \vartheta_1 = mvr \sin \vartheta_2 = M \Big|_r. \quad (5)$$

From (4) and (5) we get:

$$\vartheta_2 = \arccos \sqrt{1 - \frac{R^2 \sin^2 \vartheta_1}{r^2 \left(1 - \frac{eU}{E}\right)}} \quad (6)$$

The proton will be absorbed by the internal surface if

$$\frac{2d_i}{\cos \vartheta_2} > l(E - eU), \quad (7)$$

where $l(E')$ is the path length of a proton with energy E' in the material of the internal shell (when formulating condition (7) it was assumed that the effects of scattering due to collisions with atoms of the internal shell can be neglected). Substituting (6) into (7), we get the additional restriction on the initial direction:

$$\begin{aligned} \vartheta_1 > \vartheta_1^{**}(E) = \\ = \arccos \sqrt{1 - \frac{r^2}{R^2} \left(1 - \frac{eU}{E}\right) \left(1 - \frac{4d_i^2}{l^2(E - eU)}\right)} \end{aligned} \quad (8)$$

Let E^* be a root of the equation $l(E - eU) = 2d_i$. If $E < E^*$ and $\vartheta_1 < \vartheta_1^*(E)$ condition (7) is met automatically. However, if $E > E^*$, the restriction on the admissible angle of entrance has the form

$$\vartheta_1^{**}(E) < \vartheta < \vartheta_1^*(E). \quad (9)$$

Taking these relations into account, the total current of protons absorbed by the internal shell of the HEG is determined by the integral

$$\begin{aligned} J_{\text{abs}}^p = 8\pi^2 R^2 e \left[\int_{eU}^{E^*} \int_0^{\vartheta_1^*} \frac{dI^p}{d\omega dE} \sin \vartheta dE d\vartheta \right. \\ \left. + \int_{E^*}^{\infty} \int_{\vartheta_1^{**}}^{\vartheta_1^*} \frac{dI^p}{d\omega dE} \sin \vartheta dE d\vartheta \right], \quad (10) \end{aligned}$$

where $\frac{dI^p}{d\omega dE}$ is the differential intensity of protons of the Earth's radiation belt.

In the considered stationary mode of operation (the current J_{abs}^p is balanced by current J_1 flowing through the load) the power $W = UJ_{\text{abs}}^p(U)$ is released on the load. In this case, the specific power w of the generator, equal to the ratio of the power W to the generator's mass $M_g = 4\pi(\rho_i d_i r^2 + \rho_e d_e R^2)$, is determined by the formula:

$$w = \frac{2\pi eU}{\rho_i d_i \frac{r^2}{R^2} + \rho_e d_e} \times \left\{ \int_{eU}^{E^*} \left[1 - \sqrt{1 - \frac{r^2}{R^2} \left(1 - \frac{eU}{E} \right)} \right] \frac{dI^p}{d\omega dE} dE + \int_{E^*}^{\infty} \left[\sqrt{1 - \frac{r^2}{R^2} \left(1 - \frac{eU}{E} \right)} \left(1 - \frac{4d_i^2}{l^2(E - eU)} \right) - \sqrt{1 - \frac{r^2}{R^2} \left(1 - \frac{eU}{E} \right)} \right] \frac{dI^p}{d\omega dE} dE \right\} \quad (11)$$

In the particular case when the radii of the external and internal shells are close and the thickness of the external shell is small in comparison with the thickness of the internal shell, the formula (11) for w reduces to the simpler form

$$w = \frac{2\pi eU}{\rho_i d_i} \left\{ \int_{eU}^{E^*} \left[1 - \sqrt{\frac{eU}{E}} \right] \frac{dI^p}{d\omega dE} dE + \int_{E^*}^{\infty} \left[\sqrt{\frac{eU}{E}} + \frac{4d_i^2}{l^2(E - eU)} \left(1 - \frac{eU}{E} \right) - \sqrt{\frac{eU}{E}} \right] \frac{dI^p}{d\omega dE} dE \right\}. \quad (12)$$

From (12) follows that in this particular case the specific power of the HEG does not depend on radii of its shells and is determined by the differential intensity of the proton flux in NES, by the voltage between the shells, and also by the density and thickness of the material of the internal shell.

Using formula (12), let us calculate the specific power of a HEG whose internal shell is manufactured from an aluminum foil with a thickness of 0.5 to 8 microns. In this case the current of protons absorbed by the internal shell of the generator is determined by the flux of low-energy protons with energies $\leq 1\text{MeV}$, whose differential intensity flux near the equatorial plane is described

by the exponential function [2, 11]

$$\frac{dI^p}{d\omega dE} = \frac{I_0^p}{E_0} \exp \left[-\frac{E}{E_0} \right], \quad (13)$$

where I_0^p is the total flux of protons per unit solid angle. In order to determine the values of I_0^p at various magnetic shells, let us take advantage of the data about proton fluxes in NERB in the plane of geomagnetic equator as presented in [1] for the maximum phase of solar activity. The dependence of the energy in (13) on the McIlwain parameter L can be approximated by a power-law, $E_0 \sim L^{-3}$ [2, 11]. According to [11] E_0 is fixed for $L = 5$ to $E_0 = 120\text{keV}$. Furthermore, the dependence of the path length of a proton in the internal shell on its energy E' is described by an analytical approximation of the function $l(E')$ as presented in [9].

4 Results for the HEG specific power inside the natural radiation belts of the Earth

From the results obtained above the specific power of a HEG located in NES with values of L ranging from $L = 2.5$ up to $L = 4.5$ shall be calculated. First of all we notice that—as can be expected from the calculations—at a fixed location in the natural ERB the specific power of the HEG with a given thickness of the internal shell varies non-monotonically as a function of the operating HEG voltage, reaching the maximum value w^{max} at a certain operating voltage of the HEG which is optimal for its given location in NES. For the considered values of the thickness of the HEG's internal shell (from 0.5 to $8\mu\text{m}$) the optimum operating voltages U^{opt} decrease monotonically with increasing parameter L (see Fig. 2). The calculated L dependences of the maximum specific power of the HEG for different values of the thickness of the generator's internal shell are presented in Fig. 3. As is seen from this figure, at a fixed thickness of the internal shell the power w^{max} varies non-monotonically with increasing L -parameter, reaching the largest value at a certain optimal (for the considered value of d_i) location of the HEG in natural ERB. The limiting value of the specific power of the HEG, equal to 3.3W/kg , is reached at a thickness of the internal shell of $1\mu\text{m}$ for a location of the generator on the magnetic shell $L = 3.3$.

Now let us consider the issue of possible values of the HEG specific power in NES regions where the flux of the ERB electrons is dominant. In accordance with the data of [2] these regions are located on magnetic shells

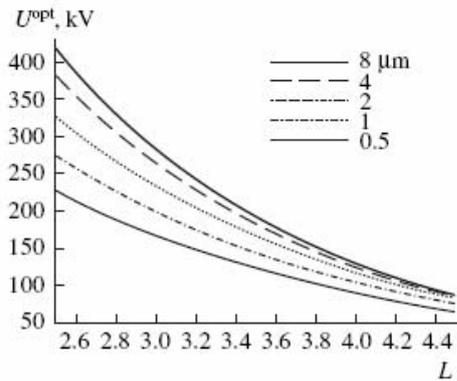


FIGURE 2.

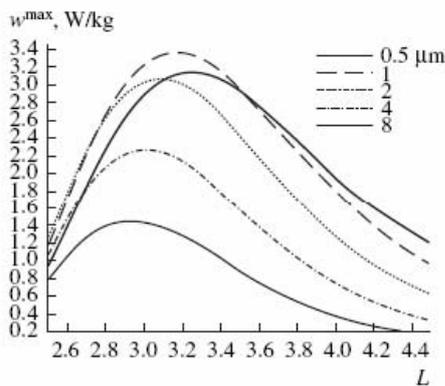


FIGURE 3.

with $L \leq 2$ and $L \geq 5$. As in the case considered above we take advantage of formula (12) to determine specific power of the HEG. One should have in mind that estimates obtained in this way are approximate, since, unlike the case of protons, the effects of scattering of electrons during their passage through generator's shells can play an important role when the HEG is charged by the flux of ERB electrons. These effects were not taken into account in derivation of formula (12).

We first make an estimate of possible values of the HEG specific power in the inner radiation belt of the Earth on magnetic shells with values of L from 1.4 to 2. We approximate the differential intensity of the electron flux by an exponential function similar to (13) with $E_0 = 200\text{keV}$ for $L = 1.4$ and $E_0 \sim L^{-2}$ for $L > 1.4$ [2]. In order to determine the values of I_0^e on various magnetic shells the data on the flux of ERB electrons in the plane of the geomagnetic equator as presented in [2] for the

phase of maximum solar activity is used.

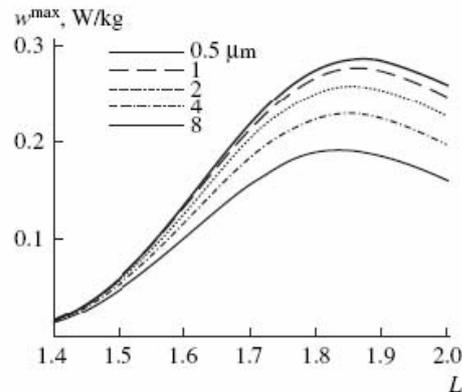


FIGURE 4.

The optimal operating voltage U^{opt} for the considered values of the thickness of the HEG internal shell (from 0.5 up to $8\mu\text{m}$) monotonically decreases with increasing L . Figure 4 presents calculated L dependences of the maximum specific powers of the HEG for various values of the thickness of the generator's internal shell. As is seen from Fig. 4, at a fixed thickness of the internal shell the power w^{max} reaches its largest value at a certain optimal (for the considered value of d_i) location of HEG in the inner electron radiation belt of the Earth. The limiting value of the specific power of the HEG, equal to 0.29 W/kg , is reached in this case at a thickness of the internal shell of $0.5\mu\text{m}$ for the generator's location on the magnetic shell $L = 1.9$.

Finally, let us estimate possible values of the specific power of the HEG in the outer electron radiation belt of the Earth. The differential intensity in this case was derived from the experimental data of [12]. It follows that the specific power w^{max} monotonically decreases with increasing thickness of the internal shell of the HEG from its maximum value equal to $2.9 \cdot 10^{-2}\text{ W/kg}$ at $d_i = 0.5\mu\text{m}$ down to $w^{\text{max}} = 1.4 \cdot 10^{-2}\text{ W/kg}$ for $d_i = 8\mu\text{m}$.

5 Discussion of the results for ERB

The results of the above estimate of possible values of the specific power of the HEG, which transforms the kinetic energy of particles of the natural radiation belts into electric power, demonstrate that the maximum values of the generator's specific power (up to 3.3 W/kg)

can be realized in the case when the generator is placed in the gap between the inner and outer radiation belts. In this region the flux of protons of the ERB substantially exceeds the flux of energetic electrons. The maximum values of the specific power are by an order of magnitude smaller if the HEG is placed in the inner electron radiation belt and by two orders of magnitude in the outer electron belt. For comparison, notice that the specific power of available systems of power supply for small satellites is within the limits of 1–3 W/kg [1] at a total power of ~50–100 W. Thus, the HEG construction under consideration turns out to be quite competitive (at least in the region where the flux of ERB protons dominates) with the power sources for small spacecraft that operate using traditional principles. Moreover, in those special cases when an output voltage of tens or even hundreds of kilovolts is needed (for example, as energy supply of ion engines or onboard accelerators of charged particles) it is preferable to use a HEG, since it produces the electric power directly at high primary voltage, which makes it unnecessary to use additional devices of voltage rise.

However, one should have in mind that the above estimates of the specific power of a high-voltage electric generator are a first step in studying the circle of science and engineering problems associated with the development of the HEG. First of all, the evaluation of possible influence of various currents due to charge leakage through the high-voltage vacuum gap on the process of the HEG charging is among these problems. Discharge currents and currents of secondary electron emission are basic currents of this type. If the HEG charging is caused by the flux of ERB protons, one can neglect the contribution of discharge currents to the process of charging, provided that their density is $\leq 10^{-8} \text{ A/m}^2$. As is shown by results from in situ experiments on charging a conducting shielded body onboard of the satellites of the *Kosmos* series [7, 8, 6], the density of discharge leakage currents at voltages below 100 kV does not exceed $\sim 10^{-9} \text{ A/m}^2$, i.e., it is smaller than the above limiting value by at least one order of magnitude. However, the necessary small values of discharge currents, $\sim 10^{-9} \text{ A/m}^2$, are achieved only in the case of pre-burning high-voltage currents with a density of $\sim 10^{-2} - 10^{-4} \text{ A/m}^2$ and a duration of tens of minutes. If the generator operates in the NES region with a predominance ERB proton flux, the current of secondary electrons with energies of ~1 eV can be shown to have a substantial influence on the process of HEG charging. These electrons are generated in the material

of the HEG external shell when the high-energy protons penetrate into the space between the shells. In this case, one should include into the HEG construction an additional grid positioned near the inner surface of the external shell to cut off the extraneous current due to the secondary electrons. This grid must be kept under a small negative potential of the order of a few volts. If the generator operates in the electron radiation belts of the Earth, the current of secondary electrons may prevent the creation of a high voltage between the HEG shells. However, one can demonstrate that due to the smallness of the coefficient of the secondary electron emission at energies typical for ERB electrons, this current can be neglected in the case under consideration.

6 Results for the specific power of the HEG in the polar region

The HEG could be put in a polar region where intense electron fluxes can be found (named fluxes of auroral electrons). The main characteristics of these fluxes are: 1) a strong time variation of total flux; 2) a strong dependence of the location of the auroral electrons on geomagnetic conditions; 3) a complex energy spectrum with several local maxima; its values can change significantly with time; 4) the energy of the first maximum is about 1 keV, while second one is located inside an interval from 15 keV to 20 keV; 5) the maximal total electron fluxes are higher than ERB fluxes; 6) the altitudes of the maximal fluxes are substantially lower than ones for other cases [10]. Experimental studies of auroral electron fluxes which took place during Defence Meteorological Satellite Program permitted to deduce an approximate expression of the auroral electron distributional functions as a combination of some Maxwellians [14]. Based on these functions we have evaluated the maximum specific power of the generator w^{max} for various values of the thickness of the generator's internal shell (Fig. 5). The obtained data shows that the specific power can exceed 100 W/kg for the thickness of 100 nm. If the thickness is less than 100 nm our evaluations prognosticate an even higher efficiency. However in this case we have to use improved methods of electron path length calculations in the material and modeling of charging processes.

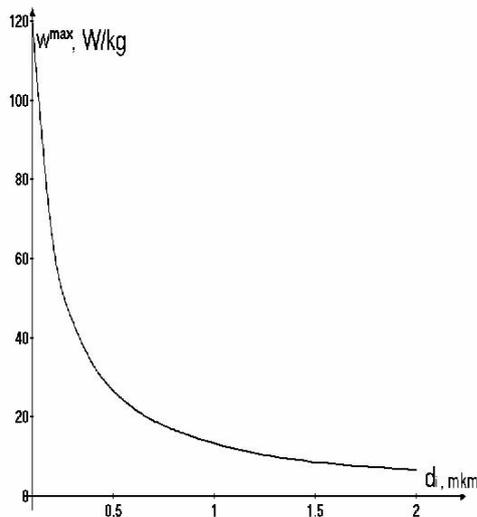


FIGURE 5.

7 Conclusions

If the HEG is placed inside the natural radiation belts of the Earth, our calculation shows that the largest value of the HEG specific power is 3.3 Wt/kg. In the inner electron radiation belt the maximum values are smaller by about one order of magnitude, in the outer electron belt by up to two orders of magnitude. If the HEG is placed in the auroral zone our estimates suggest that a specific power can exceed 100 W/kg.

Considering its weight the construction of the HEG as considered in this paper turns out to be quite competitive with power plants for small spacecraft operating on conventional principles. This in particular applies for those special cases where the power supply should provide high output voltage of tens or even hundreds of kilovolts.

Together with further investigations on the influence of leakage currents through the high-voltage vacuum gap on functioning of HEG, problems like the maintenance of the stiffness and the stability of the generator construction or the calculations of its strength characteristics should be considered.

In conclusion, we emphasize that the Earth is not the only planet in the solar system with radiation belts. Powerful radiations belts (far exceeding in their particle density the Earth's radiation belts) have been found in the vicinity of Jupiter. Furthermore, the existence of radiation belts is established for Saturn and Uranus. At

the same time, at large distances from the Sun the specific power of conventional electric power sources based on converting the energy of solar electromagnetic radiation into electric power becomes insignificant. Therefore, one cannot exclude that the principle of getting electric power in space as proposed here may turn out to be especially efficient for power supply of spacecraft in the vicinity of giant planets of the Solar system.

References

- [1] Y. Afagonov, Y. Voita, S. Gruzkov, and S. Shpakov. *Sistemy elektropitaniya malykh sputnikov. Obzor informatsii (Systems of Power Supply for Small Satellites: A Review)*. IKI RAN, Moscow, 1997.
- [2] S. Avakyan, A. Vdovin, and V. Pustarnakov. *Ioniziruyushchie i pronikayushchie izlucheniya v okolozemnom kosmicheskom prostranstve. Spravochnik (Ionizing and Penetrating Radiations in Near-Earth Space: A Handbook)*. Gidrometeoizdat, St. Petersburg, 1994.
- [3] A. Bolonkin. Space electric generator run by solar wind. In *World Space Congress: 43rd Int. Astronaut. Congr. (IAF) and 29th Plen. Meet. Comm. Space Res. COSPAR*, pages 182–183, 1992.
- [4] S. Deforest. Spacecraft charging at synchronous orbit. *J. Geophys. Res.*, 77:651–659, 1972.
- [5] E. K. Kolesnikov. Evaluation of the Specific Power of a High-Voltage Generator Based on the Use of Energy of the Earth's Radiation Belts. *Kosm. Issled.*, 44:486–492, 2006.
- [6] E. Kovalev, E. Molchanov, V. Lebedev, T. Ryabova, and I. Pekhterev. Investigation of electrostatic charging of a conducting screened body on KOSMOS-936 satellite. *Kosm. Issled.*, 25:585–591, 1987.
- [7] E. Kovalev, E. Molchanov, and T. Ryabova. Investigation of basic properties of electrostatic shield of the KOSMOS-605 satellite. *Kosm. Issled.*, 13:771–777, 1975.
- [8] E. Kovalev, E. Molchanov, and T. Ryabova. Investigation of basic properties of electrostatic shield of the KOSMOS-605 satellite. Part II. *Kosm. Issled.*, 14:126–132, 1976.

- [9] O. Nemets and Y. Gofman. *Spravochnik po yadernoi fizike (Handbook on Nuclear Physics)*. Naukova dumka, Kiev, 1975.
- [10] J. Tien, S. Akasofu, and S. Chapman. *Solar-Terrestrial Physics*. Clarendon Press, Oxford, 1972.
- [11] B. Tverskoi. *Dinamika radiatsionnykh poyasov Zemli (Dynamics of the Earth's Radiation Belts)*. Nauka, Moscow, 1986.
- [12] P. Vakulov, V. Vorobev, S. Kuznetsov, I. Logachev, I. Savenko, and V. Stolpovskii. Investigation of captured radiation from KOSMOS-426. II-structure of the radiation belts. *Kosm. Issled.*, 13:945–948, 1975.
- [13] M. Vignoli, W. Miller, and M. Matteoni. Power generation with electrodynamic tethers. In *Tethers in Space: Proc. Int. Conf.*, pages 473–481, San Diego, 1987.
- [14] H. Yeh and M. Gussenhoven. The statistical electron. Environment for defense meteorological satellite program eclipse charging. *J. Geophys. Res.*, 92(A7):1105–7715, 1987.

Panel 3: "Driving Evolution" – A paradigm shift in the balance between nature and technique? Consequences and opportunities for space?

DRIVEN by new powerful techniques discovered during the second half of the 20th century, biology is frequently considered as the science of the 21st century. Since years, media abounds with examples describing the dangers and consequences of the battle between nature's evolution and frantic technological progress. As a result, the worry about the ability of nature and technology to adapt to each other is spreading in people's minds.

On one hand, the powerful explanatory *evolution's* theory has become the central organizing principle of modern biology, which itself plays an increasing role in inspiring advanced technologies, in particular a hybrid discipline generally called "biomimetics". Moreover, *evolution* is an adaptation at various scales of the individuals to environmental conditions, whose alteration is partly due to technological progresses. In that sense, technology constitutes an intrinsic parameter of *evolution*. On the other hand, Heidegger wrote that *technique* has a fate in itself: it is not just a neutral external mean whose value is given by the use we make of it, but also as a way individuals use to express their "destinal" nature, which *technique* inherits and which becomes its own finality. To this extent, the nature of human being - a result of *evolution* - is a constitutive parameter of *technique*.

In a few words, evolutions of nature and technology are closely cross-linked. From molecular mechanisms and carbon nanotubes to multiscale surface interaction and locomotion systems, technology uses analogies with nature's mechanisms at various scales, and by facilitating the fulfillment of human's needs, smart environment using last advances in e.g. robotics constitute another bridge between *evolution's* and nature's evolutions.

What is in there for space? What will be the role of "dark, cold, empty, harsh and hostile" space in an evolution that is "life-centred", that places the very nature of life and humans into the core of its discussion? The process can be expected to affect and question basic principles, from the value of natural systems, biodiversity and the right balance between the increasingly

indistinguishable "natural" and "artificial". What could be the role of space (physically, as well as a tool)? So far, exploiting particular space conditions has led to applications which have improved some aspects of life on Earth. Whereas research is exploring the use of the very specific properties of space environments, no commercial or larger scale in-space production has yet been identified. Could "driving evolution" change this equation? Are we looking into biological systems "engineered for space" or rather "engineered in space"? How could this affect human presence in space?

This panel will not answer these questions, but with these in mind, it addresses some first timid steps, e.g. using bio-inspired techniques to attach and detach: could these be used to connect spacecrafts, units to spacecrafts or further in the future to small bodies like asteroids?

For centuries, human brains have been dreaming about space: how close are we now to make some room in space for humans?

François Nuyttens
Panel coordinator



Acta Futura 3 (2009) 91-95

DOI: 10.2420/AF03.2008.74

**Acta
Futura**

Playing with forces and interactions to manipulate single molecules

BOUZID MENAA, ANNE-SOPHIE DUWEZ*

Department of Chemistry, University of Liège, B6a Sart-Tilman, B-4000 Liège, Belgium

Abstract. Molecular manufacturing is a technology that will allow us to assemble molecular machines and build complex objects atom by atom. The use of scanning probe microscopy-based techniques to manipulate single molecules [10], to detect binding processes [20, 18, 21, 3, 5, 23, 22, 6, 17, 12, 14, 7, 1, 13], or to deliver molecules in a precisely controlled manner to a specific target [11, 19] represents a significant step in that direction. It requires the controlled formation and breaking of individual bonds. Here we show that the atomic force microscope (AFM) can deliver and immobilize single molecules, one at a time, on a surface. Reactive polymer molecules, attached at one end to an AFM tip, are brought into contact with a substrate to which they become linked by a chemical reaction. When the AFM tip is pulled away from the surface, the resulting mechanical force causes the weakest bond the one between the tip and polymer to break. This process transfers the polymer molecule to the substrate. We also show examples of the use of those AFM tips bearing reactive polymers for molecular recognition applications. We have covalently attached proteins or ligands to those tips to obtain probes sensitive to specific molecular interactions. We can imagine that the functional principles and concepts found in molecule manipulation by AFM, i.e. playing

with mechanical forces, with strong and weak complementary interactions, could be implemented to attach and detach units in the space field.

1 Introduction

Since its invention in the late 1980s, the atomic force microscope (AFM) has increasingly been used for the visualization of molecular systems and complex biological structures. The tight attachment of (bio)molecules onto AFM tips has opened up the exciting possibility of detecting binding processes at the single molecule level [20, 18, 21, 3, 5, 23, 22, 6, 17, 12, 14, 7, 1, 13]. Making and breaking bonds in a controlled way is strongly dependent on the design and accurate functionalization of the probes. AFM tips functionalized with end-grafted molecules offer the prospect of delivering individual molecules in a single-molecule force spectroscopy experiment if the bond anchoring the molecule to the tip is weaker than the one to be established with the surface. In order to achieve this goal, we developed a strategy based on (1) the grafting of macromolecules bearing reactive groups onto the AFM tip and (2) their selective transfer, via a chemical reaction, to a substrate where complementary moieties are present [4]. AFM tips with reactive macromolecules can also be used for recognition experiments on surfaces where receptors are grafted. We

*Corresponding author. E-mail address: asduwez@ulg.ac.be

demonstrate that such tips are a robust basis for single molecule recognition between host-guest systems [2].

2 Single molecule delivery experiments

Gold-coated AFM tips were modified by electrografting poly-N-succinimidyl-acrylate (PNSA), according to [16]. This electro-initiated polymerization is a convenient way to fabricate polymer brushes with a moderate grafting density and results in the direct chemisorption of the polymer onto the tip surface. The cantilever to be grafted is simply dipped into the monomer solution and is selectively polarized on the tip side in the cathodic range until the so-called 'grafting peak' is observed, using a classical three-electrode setup [16, 9]. It is an electro-initiated process that requires the presence of a few charges only for the grafting step. The chain propagation that follows is a chemical process that does not need current for being sustained. The polymer selected for grafting is PNSA. The choice is based on the high room temperature reactivity of the activated esters along the backbone, paving the way to further easy coupling reaction with nucleophilic compounds, both in water and in organic solvents [15]. Substrates where complementary moieties are present were also prepared. As the activated esters along the polymer backbone can easily react, at room temperature, with amino-derivatives, aminopropyltri-methoxysilane was grafted to silicon substrates to obtain an amino-terminated surface. See [8] for more details. In a DMF solution containing 4 dimethylaminopyridine (DMAP, a catalyst), the functionalized AFM tip was slowly brought into contact with the surface. The chemical reaction between the PNSA activated esters and the amino groups of the substrate forms amide bonds and covalently links polymer chains to the substrate. Upon retraction of the tip, single chains are stretched until a bond breaks. The Au(tip)-C(polymer) bond is the weakest link in the system and the most likely candidate for breaking. Upon cleavage, the polymer chain remains covalently attached to the substrate (Figure 1). The deposited chains are reactive and can be easily post-functionalized by a wide range of nucleophilic compounds. The stretching of the polymer chains and the mechanical breaking of the Au-C bond, and thus the successful delivery is monitored through force-distance curves.

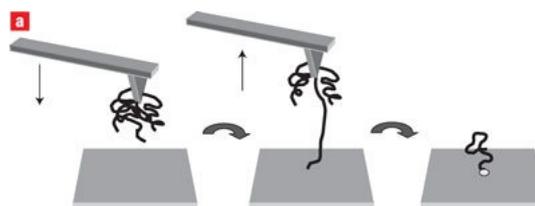


FIGURE 1. Molecule by molecule delivery process: PNSA chains grafted to the Au-coated tip are brought into contact with an amino-modified silicon substrate to which they can become linked through the formation of an amide bond, which covalently links the chain to the substrate. When the tip is pulled away from the surface, the resulting mechanical force causes the weakest bond—the one between the tip and polymer—to break. Adapted from [4].

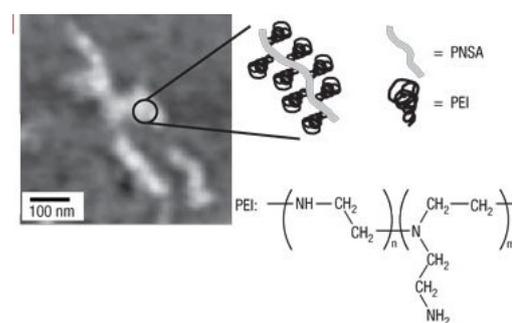


FIGURE 2. AFM topography image obtained in air in the area where 4 PNSA chains were deposited one at a time. The original chains were decorated by a branched polyethyleimine (PEI), rinsed with DMF, and imaged before the residual film of DMF was completely evaporated. The decorated molecules appear in an extended shape. Maximum vertical height: 4 nm. Adapted from [4].

3 Molecular recognition experiments

Recognition experiments between two complementary host-guest molecules can be realized through force spectroscopy measurements between one partner grafted onto the AFM tip and the other partner grafted onto a surface (Figure 3). In order to carry out reliably force spectroscopy measurements, a series of experimental requirements are to be met. First, the binding force of the chemical or biological species onto the tip has to be stronger than the investigated interaction strength. Second, the presence of a long-chain spacer linking the species to the tip is required to avoid the interference of non-specific adhesion forces, which otherwise dominate when tip and surface are close to each other. PNSA was grafted from the tip [9]. As a result, chemisorption

ensures strong bonding to the tip, polymerization provides a long-chain spacer, which enables free orientation of the species, while the presence of N-succinimidyl moieties allows for further functionalization since N-succinimidyl activated esters react easily at room temperature with amine functions of whatever protein, giving rise to amide bonds. We recently showed that this polymer could be grafted in an isolated chain regime, resulting in the grafting of only one linker at the tip apex, paving the way to its use as a generic platform to probe molecular interactions [9].

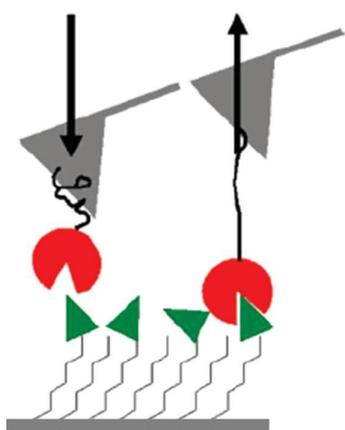


FIGURE 3. Schematic representation of a molecular recognition event investigated by single molecule force spectroscopy experiment.

One system consists of the P.69 pertactin protein (also known as antigen 69K) grafted onto a surface recognizing its complementary antibody, anti-69K. Antigen 69K is a compact globular protein located in the outer membrane of *Bordetella pertussis* and is a key component of acellular vaccines against whooping cough. The second system is based on the surface-grafted antibody 286F7, which is directed toward the thyroid stimulating hormone (TSH, also known as thyrotropin). TSH is produced by and stored in the pituitary gland, which is located beneath the brain, and its detection is used for the diagnosis of problems affecting the thyroid gland. Force-distance curves were recorded for the two ligand-receptor pairs. The curves exhibit rupture peaks far away from the adhesive force region, which reflect the antigen-antibody bond rupture [20, 18, 21, 3, 5, 23, 22, 6, 17, 12, 14, 7, 1, 13]. The value of the force at this peak is interpreted as the “unbinding force” of the antigen antibody complex, which dissociates when the applied pulling force exceeds the bonding strength [20, 18, 21, 3, 5, 23, 22, 6, 17, 12, 14, 7, 1, 13].

The distance at which the bond rupture occurs depends on the position of the protein on the PNSA linker. For a given tip, it should thus always occur at the same distance. The measurement of a large number (more than 250) of force curves, all recorded at the same loading rate allowed us to construct force histograms. They display a close-to-Gaussian profile, whose maximum provides the most probable antigen-antibody rupture force. The force distribution recorded for the 69K Ag/anti-69K Ab couple gives a most probable unbinding force of 256 pN at a loading rate of 100 nN s^{-1} . Likewise, the most probable unbinding force between 286F7 Ab and TSH is 76 pN [2]. In order to confirm that the measured forces do correspond to the rupture of the antigen-antibody complex, the force-distance curves were compared with those recorded between tip and surface systems not expected to interact with each other. In particular, force-distance curves were recorded between a 69K Ag grafted surface and an unmodified PNSA tip, between a methyl-terminated, self-assembled monolayer surface and a PNSA tip modified with anti-69K Ab, and between a 69K Ag functionalized surface and a PNSA tip modified with 286F7 Ab. In that case, the retraction curve is characterized by a flat line where only small adhesive forces sometimes appear, testifying for the absence of any specific interaction established between these incompatible pairs. Based on all these arguments, it can thus be concluded that the rupture peaks observed are indeed due to the specific interaction between the complementary antigens and antibodies.

4 Conclusions

Molecular manufacturing is a technology that will allow us to assemble molecular machines and build complex objects atom by atom. This ultimate goal may be a long way off, but recent developments in single molecule manipulation with AFM-based techniques move us closer to the idea of positioning individual atoms and molecules. Richard Feynman said in his famous talk in 1959: “The principles of physics, as far as I can see, do not speak against the possibility of maneuvering things atom by atom.” Adding positional control to chemical reactions represents a significant step towards molecular manufacturing. Our current research activities focus on the development of AFM techniques to make molecules go where we want, put them where we want, and make them react as we want. The controlled manipulation of molecules requires the forma-

tion and breaking of targeted individual bonds. We can imagine that the functional principles and concepts found in molecule manipulation by AFM, i.e. playing with mechanical forces, with strong and weak complementary interactions, could be implemented to attach and detach units to spacecrafts.

References

- [1] C. Bustamante, Y. R. Chemla, N. R. Forde, and D. Izhaky. Mechanical processes in biochemistry. *Annu. Rev. Biochem.*, 73:705–748, 2004.
- [2] F. Cecchet, A.-S. Duwez, S. Gabriel, C. Jerome, R. Jerome, K. Glinel, S. Demoustier-Champagne, A. M. Jonas, and B. Nysten. Atomic force microscopy investigation of the morphology and the biological activity of protein-modified surfaces for bio- and immunosensors. *Anal. Chem.*, 79:6488–6495, 2007.
- [3] M. Conti, G. Falini, and B. Samorì. How strong is the coordination bond between a histidine tag and Ni - Nitriolotriacetate? an experiment of mechanochemistry on single molecules. *Angew. Chem. Int. Ed.*, 394:215–218, 2000.
- [4] A.-S. Duwez, S. Cuenot, C. Jerome, S. Gabriel, R. Jerome, S. Rapino, and F. Zerbetto. Mechanochemistry: targeted delivery of single molecules. *Nature Nanotechnol.*, 1:122–125, 2006.
- [5] R. Eckel, R. Ros, B. Decker, J. Mattay, and D. Anselmetti. Supramolecular chemistry at the single-molecule level. *Angew. Chem. Int. Ed.*, 44:484–488, 2005.
- [6] R. Eckel, S. D. Wilking, A. Becker, N. Sewald, R. Ros, and D. Anselmetti. Single-molecule experiments in synthetic biology: an approach to the affinity ranking of DNA-binding peptides. *Angew. Chem. Int. Ed.*, 44:3921–3924, 2005.
- [7] E. Evans. Probing the relation between force–lifetime–and chemistry in single molecular bonds. *Annu. Rev. Biophys. Biomol. Struct.*, 30:105–128, 2001.
- [8] C. A. Fustin, G. Glasser, H. W. Spiess, and U. Jonas. Parameters influencing the templated growth of colloidal crystals on chemically patterned surfaces. *Langmuir*, 20:9114–9123, 2004.
- [9] S. Gabriel, C. Jerome, R. Jerome, C.-A. Fustin, A. Pallandre, J. Plain, A. M. Jonas, and A.-S. Duwez. One-step polymer grafting from silicon nitride SPM probes: From isolated chains to brush regime. *J. Am. Chem. Soc.*, 129:8410–8411, 2007.
- [10] J. K. Gimzewski and C. Joachim. Nanoscale science of single molecules using local probes. *Science*, 283:1683–1688, 1999.
- [11] D. S. Ginger, H. Zhang, and C. A. Mirkin. The evolution of dip-pen nanolithography. *Angew. Chem. Int. Ed.*, 43:30–45, 2004.
- [12] H. H. Clausen-Schaumann, M. Seitz, R. Krautbauer, and H. E. Gaub. Force spectroscopy with single bio-molecules. *Current Opinion Chem. Biol.*, 44:524–530, 2000.
- [13] P. Hinterdorfer and Y. F. Dufrene. Detection and localization of single molecular recognition events using atomic force microscopy. *Nature Methods*, 3:347–355, 2006.
- [14] A. Janshoff, M. Neitzert, Y. Oberdoerfer, and H. Fuchs. Force spectroscopy of molecular systems—single molecule spectroscopy of polymers and biomolecules. *Angew. Chem. Int. Ed.*, 394:3212–3237, 2000.
- [15] C. Jerome, S. Gabriel, S. Voccia, C. Detrembleur, M. Ignatova, R. Gouttebaron, and R. Jerome. Preparation of reactive surfaces by electrografting. *Chem. Commun.*, pages 2500–2501, 2003.
- [16] C. Jerome, N. Willet, R. Jerome, and A.-S. Duwez. Electrografting of polymers onto AFM tips: A novel approach for chemical force microscopy and force spectroscopy. *Chem. Phys. Chem.*, 5:147–149, 2004.
- [17] F. R. Kersey, W. C. Yount, and S. L. Craig. Single-molecule force spectroscopy of bimolecular reactions: system homology in the mechanical activation of ligand substitution reactions. *J. Am. Chem. Soc.*, 128:3886–3887, 2006.
- [18] L. A. Lee, G. U. Chrisey and R. J. Colton. Direct measurement of the forces between complementary strands of DNA. *Science*, 266:771–773, 1994.
- [19] J. Loos. The art of SPM: scanning probe microscopy in materials science. *Adv. Mater.*, 17:1821–1833, 2005.

- [20] V. T. Moy, E.-L. Florin, and H. E. Gaub. Adhesion forces between individual ligand-receptor pairs. *Science*, 266:257–259, 1994.
- [21] S. S. Wong, E. Joselevich, A. T. Woolley, C. I. Cheung, and C. M. Lieber. Covalently functionalized nanotubes as nanometre-sized probes in chemistry and biology. *Nature*, 394:52–55, 1998.
- [22] Q. Zhang, J. Jaroniec, G. Lee, and P. E. Marszalek. Direct detection of inter-residue hydrogen bonds in polysaccharides by single-molecule force spectroscopy. *Angew. Chem. Int. Ed.*, 44:2723–2727, 2005.
- [23] S. Zou, H. Schönherr, and G. J. Vancso. Stretching and rupturing individual supramolecular polymer chains by AFM. *Angew. Chem. Int. Ed.*, 44:956–959, 2005.

Panel 4: “Discovering Natural Paths”

SPACE is an apparently quiet place, it may resemble a desert where, at a first glance, not much is happening and not many resources are available. Yet, looking closely, we understand that even though deserts are places we humans are not used to living in, they have many things to offer to those who accept the challenge. Certain ants, for example, are able to plan their route finding, with little or no visual reference, first to food and then the way back to their nests (a find-the-needle-in-the-haystack situation). How do they do it? Apparently, there is enough information available to them that we just cannot perceive in the first place. Does the same apply to the space environment? Do we “just” need to look more closely to discover ways and solutions that are out there already and we just do not see? Past research has shown that rather than fighting against nature’s laws, it is sometimes possible and much easier to exploit in an advantageous and smart way. What does it mean for space?

Every month the Earth gets bombarded with rocks coming from Mars, from asteroids and from comets. If we could fully understand the natural path they find to reach Earth, could we improve our sample-return missions?

When a satellite enters the space plasma it naturally gets a negative charge. System engineers often try to fight this by using neutralizing devices in order to minimise the disturbances on the spacecraft payloads. Is the electrostatic charge a disturbance or can it be considered as an actuating force? Do we really need to fight it?

Locating a spacecraft position in the interplanetary medium is a difficult task. We use sophisticated operational procedures unnecessary if spacecraft could exploit information that may already be throughout the solar system. Just like the little desert ants are always able to find their way home. After all, from pulsars to planetary ephemerides, the space is full of precise signals that a clever ant would use. Are our spacecraft that smart?

Formation flying missions constantly fight against Kepler laws as we try to force the satellites in an unnatural dynamic. But again, should we fight nature or is the solution of our problems already embedded into its laws?

The goal of this panel is to have a discussion about/on new solutions for spacecraft navigation and control based on the exploitation of often overlooked natural features.

Dario Izzo
Panel coordinator



Electrostatic Force for Swarm Navigation and Reconfiguration

LORENZO PETTAZZI^{*†}, HANS KRÜGER[‡], STEPHAN THEIL[‡], DARIO IZZO[§][†]ZARM, University of Bremen, Germany[‡]DLR Institute of Space Systems, Bremen, Germany[§]ESA-ACT, Noordwijk, The Netherlands

Abstract. In this work the concept of a swarm of satellites controlled by a hybrid thrusting electrostatic actuation system is assessed. On one side the propulsion system is investigated. First from the model of the interaction between the space plasma and a charged spacecraft a set of requirements is derived. This allows to define an actuation system for charge control. Then the applicability of the electrostatic actuation for formation keeping and reconfiguration of swarms of satellites is assessed. In particular this work aims at demonstrating that the electrostatic actuation can be exploited in a decentralized control scheme to trigger high fuel savings in reconfiguration maneuvers of swarms of satellites. To this end a novel charging strategy has been developed. The resulting system has been tested under different possible simulations and it has shown good performances in terms of reduction of the fuel expenditure for the whole swarm.

1 Introduction

Recently it was identified the possibility to perform active electric charge control on spacecraft in order to achieve a certain relative dynamic exploiting the inter-satellite electrostatic forces [3]. On the basis of this re-

sult in [6] a charging control law able to stabilize the motion of a formation of satellites only by means of the electrostatic force was developed. Only in a recent paper [4] a hybrid approach has been considered where a chief satellite can control its position by thrusters and in the meantime exploit the charge control to deploy a group of deputy satellites. However, being the electrostatic forces internal, it is not possible to use them to control the position of the formation center. For this and some other reasons, it is quite likely that if electrostatic forces will ever be used to control the relative motion or positioning of orbiting satellites, these will be coupled with conventional thrusting propulsion systems, able to control the absolute position of the satellites. In the present work we discuss the possibility to use such a hybrid actuation system in connection with decentralized path planning algorithms to improve the fuel efficiency of acquisition and reconfiguration maneuvers of swarms of satellites. Following this idea the problem is not anymore to find a formation that can be stabilized or a maneuver that can be steered only relying upon the inter-spacecraft electrostatic interaction. Rather, the proposed approach tries, for a general maneuver, to find the instantaneous spacecraft charge levels that minimize the residual force to be provided by the thrusting actuation system. To this end our investigation will be divided in two parts. In the first part the spacecraft interaction with the surrounding

^{*}Corresponding author. E-mail address: lorenzo.pettazzi@mbda-systems.de

plasma is investigated to define requirements and evaluate the charge actuation system. Then we introduce a novel charge control feedback to derive, from the output of the path planning law, a course of desired charges for the spacecraft belonging to the swarm. Finally, the performance of the charge control feedback is illustrated by means of simulations.

2 Hybrid propulsion system definition and evaluation

2.1 Model of space charging

In several works [3, 4, 5, 6] the geostationary (GEO) environment was identified as a suitable environment for the application of inter-spacecraft electrostatic forces for formation control. The space plasma in GEO consists mainly of electrons and protons. In this region of space several natural phenomena exist, which tend to change the electrical charge of the spacecraft. These are mainly represented by the solar radiation and complex interactions between the Earth's magnetosphere, the local space plasma and the solar wind. This causes the plasma environment to change over time and space and leads to varying fluxes of charged particles to and from the satellite. The incoming fluxes are the primary electron I_e and the primary proton I_i current. Furthermore there is the secondary electron emission from the spacecraft. This is mainly caused by the impact of primary electrons, while a smaller portion is caused by the impact of primary ions and backscattered secondary electrons. Another important flux is the emission of photoelectrons I_{photo} during sunlit periods.

In order to define the requirements of the electrostatic actuation system, a model of the natural space charging process has been derived and is widely discussed in [5]. The model will be explained briefly here. The spacecraft is assumed to have a spherical shape, to be perfectly conductive and to have a uniform charge distribution. Neglecting the secondary and backscattered electrons, the sum of all natural currents I_{tot} amounts to:

$$I_{tot} = I_e + I_i + I_{photo}. \quad (1)$$

If the total current I_{tot} is different from zero, the spacecraft will change its charge and thus its voltage V_{SC} with respect to the plasma electric potential by:

$$\frac{dV_{SC}}{dt} = \frac{I_{tot}}{C_{SC}} \quad (2)$$

where C_{SC} is the capacitance of the spacecraft. Based on this model, voltage-current characteristics can be calculated. The characteristics give an overview of the total current, the spacecraft receives, when it is charged at different voltage levels. In Figure 1 an example for such a V-I diagram is given.

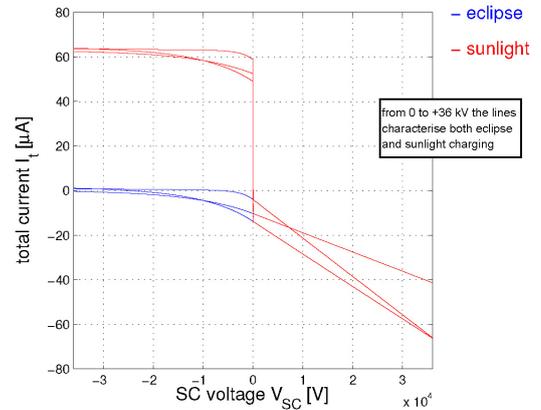


FIGURE 1. Total current in GEO for different spacecraft potentials, in different plasma environments, during sunlit and eclipse phases.

2.2 Requirements

The exploitation of inter-satellite electrostatic forces requires the control of each spacecraft's total current. Therefore, the sum of all natural fluxes has to be equalized and an additional current dependent on the desired charge state must be provided. With the utilization of a charge actuation system Eq.(1) extends to:

$$I_{tot} = I_e + I_i + I_{photo} + I_{control}. \quad (3)$$

For being able to control the spacecraft charge regardless of its own charging state, several worst case scenarios have been simulated to determine the maximum possible natural currents I_e , I_i and I_{photo} . The highest desired spacecraft charges have been set to $\pm 2 \mu C$, which translates into a maximum potential of ± 36 kV. Figure 1 displays the results, showing the highest expected total natural currents to be at $\pm 65 \mu A$. While it is necessary to compensate for the natural currents, the control current should be as low as possible to keep the mass flow at small numbers. On the other hand the spacecraft shall charge in a given time span. Considering that, the currents demanded from the actuation system have been set to $\pm 73 \mu A$. Assuming the satellite being at a certain

potential, charges with an opposite sign must be accelerated to be able to actually leave the spacecraft. Thus the actuator must be able to provide an acceleration voltage of ± 36 kV. The requirements of the charge ejection system are summarized in Table 1.

As the spacecraft shall be charged in negative and positive direction the requirements are the same for both the electron and the ion ejection system.

<ul style="list-style-type: none"> - range of acceleration voltage $V_{acc} = \pm 36$ kV - independent adjustability of $I_{control}$ and V_{acc} - fast adjustability of $I_{control}$ and V_{acc} in less than 1 s - provision of emission currents $I_{control}$ in the order of ± 73 μA

TABLE 1. Requirements of the charge actuation system [5].

2.3 Definition of the actuator

In order to assess whether the requirements defined in the previous section can be fulfilled, we propose here a candidate system to be used for charge actuation. Furthermore an algorithm for the operation of this actuator is proposed.

After a survey of available charge mitigation techniques and μ N ion thrusters in [5] it turned out, that the best suitable actuator for a charge control system is the Radio Frequency Ion Thruster (RIT), as it can be operated with a high bandwidth of emission currents and acceleration voltages and is capable to adjust these two parameters independently. The RIT will be used for the emission of positive charges. With the present technology the emission has a lower boundary in the order of 250 μ A for the RIT-4 and - possibly available in future - 100 μ A for the RIT-1. An electron gun has been selected for the emission of negative charges.

The approach presented here to operate the charge actuator builds on existing technology and uses the simultaneous emission of two currents, which differ in magnitude and sign. They comprise the charging current I_{ch} and the stabilizing current I_{st} . To charge the spacecraft, particles must be emitted such that $\frac{dV_{SC}}{dt}$ leads V_{SC} in the direction of the desired voltage. In this approach I_{ch} is emitted with an acceleration voltage V_{acc} equal to V_{des} .

E.g. to change the spacecraft potential from zero to a negative value, I_{ch} will consist of positive ions. As soon

as the satellite reaches a negative potential, the ions are attracted by it. At $|V_{SC}| < |V_{acc}|$, the particles can escape. In the moment V_{SC} assumes V_{acc} the expelled charges can not leave the spacecraft anymore and therefore cease to further charge it. Thus the desired voltage can be achieved by setting V_{acc} to the value of V_{des} .

If, due to external events like sunlight, the spacecraft would float to a positive potential, the particles of I_{ch} can escape again, bringing the satellite back to V_{des} . This is true as long as the charging current outnumbers any natural current. For positive V_{des} the same method is used, but changing the flux from positive ions to electrons.

This approach lacks the ability to charge the spacecraft below the floating potential V_f . This is the potential, the satellite would assume without any artificial current. In general, the desired voltage can not be achieved if:

$$\begin{aligned} |V_{des}| &< |V_f| \\ \text{sign}(V_{des}) &= \text{sign}(V_f) \end{aligned} \quad (4)$$

This is extensively explained in [5]. To solve that problem, the stabilizing current I_{st} is introduced. It is simultaneously operated, possessing the opposite sign of I_{ch} . While the charging current is charging the spacecraft in direction from zero to the desired voltage, I_{st} pulls it to a zero potential. Considering the application of charge control the balance in Eq. 3 changes now to:

$$I_{tot} = I_e + I_i + I_{photo} + I_{ch} + I_{st} \quad (5)$$

where $I_{control}$ is replaced by $I_{ch} + I_{st}$.

With the control scheme introduced in [5] the signs and magnitudes for both currents are set in such a way, that I_{tot} of Eq. (5) is zero when V_{SC} reaches V_{des} . Furthermore a V-I characteristic of the spacecraft-plasma system is created, which is stable around V_{des} . Thus the satellite is kept at its desired voltage level, even if the environmental parameters are changing. This is valid also for the worst case environmental conditions expected at GEO.

2.4 Performance evaluation

In this section the performance and applicability of the charge actuation system shall be evaluated in terms of the duration of the charging process, the accompanying residual forces and the mass flow. The time performance of the charging process has been evaluated by simulating the cases, where the longest charging times are expected. Therefore the control system is ordered

to switch between the given limits of desired voltages from -36 kV to $+36$ kV and vice versa. The simulations take place in both sunlit and eclipse phases, experiencing worst case conditions of the space plasma, where the natural currents have their highest magnitudes. In Figure 2 the response of the system under various environmental conditions is shown. It can be seen that the highest charging time is not greater than 100 ms.

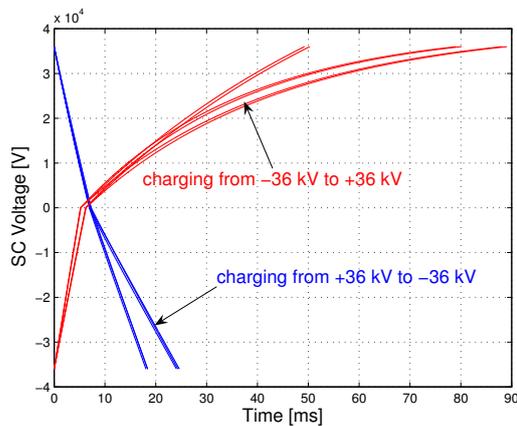


FIGURE 2. Charging durations for different worst case scenarios [5].

Besides the electrostatic inter-spacecraft forces, the emission of the control currents creates a repulsion force, which is considered here as completely residual. For the proposed devices the emitted species are electrons as negative charges and Xenon ions or protons as positive charges. For the evaluation the acceleration voltage of the ion thruster and the effect of the spacecraft potential field have been considered. The residual forces have been calculated for the case of formation keeping, where the spacecraft charges are expected to vary in small ranges and for the case of an acquisition maneuver with frequently changing V_{des} . In Table 2 the worst case residual forces due to emission of protons are given for a charge actuation system with the RIT-1 engine using hydrogen as ion delivering gas. The forces caused by electrons can be neglected. An extensive explanation can be found in [5].

To prove the applicability of electrostatic actuation, the mass flow needed for charge emission shall be compared with the mass flow needed by the same actuator to generate the same amount of force, but working as a conventional ion thruster. It turned out, that for hydrogen usage, the electrostatic actuation is in advance for

	worst case formation keeping V_{des} is varying in small range
H^+	$\ll 1 \mu N$ constant (I_{ch}) -
	worst case acquisition V_{des} may change rapidly
H^+	$5 \mu N$ short period (I_{ch}) $2.7 \mu N$ constant (I_{st})

TABLE 2. Worst case residual forces for RIT-1 using hydrogen as ion delivering gas.

forces higher than $3.5 \mu N$ (RIT-1) and $7 \mu N$ (RIT-4). The case of hydrogen as ion delivering gas is displayed in Figure 2.

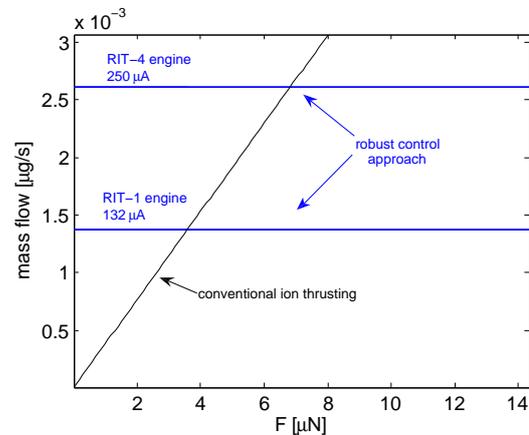


FIGURE 3. Comparison of mass flows between electrostatic actuation and ion thrusting.

The evaluation shows, that all requirements of Table 1 are fulfilled and a fast and robust control of the individual spacecraft charges can be provided.

3 Charging strategy definition

Let us consider a swarm of N of satellites having mass m_i flying in a tight formation around a circular orbit with orbital angular velocity ω . Let $\mathbf{r}_i = [x_i, y_i, z_i]^T$ be the relative position of the i -th spacecraft with respect to the center of the swarm projected onto the relevant Local-Horizontal-Local-Vertical (LHLV) frame with the x axis pointing radially away from the center of the Earth, and with the z axis pointing in the direction normal to the reference orbit plane. Then the equa-

tion of motion of the i -th spacecraft belonging to the swarm can be approximated by the Clohessy-Wiltshire equations

$$\ddot{\mathbf{r}}_i + \mathbf{D}\dot{\mathbf{r}}_i + \mathbf{K}\mathbf{r}_i = \mathbf{u}_{des_i}$$

where:

$$\mathbf{D} = \begin{bmatrix} 0 & -2\omega & 0 \\ 2\omega & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \mathbf{K} = \begin{bmatrix} -3\omega^2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \omega^2 \end{bmatrix}$$

and where, under the hypothesis of perfect actuation, \mathbf{u}_{des_i} is the acceleration command issued by the on-board path-planning system. We assume that each spacecraft is able to exploit the thrusting actuation system to generate accelerations \mathbf{u}_{t_i} in any direction and that the net currents I_i flowing from the spacecraft can be actively modulated to control the electrical charge q_i of each satellite. Then

$$\begin{aligned} \mathbf{u}_{des_i} &= \\ \mathbf{u}_{e_i} + \mathbf{u}_{t_i} &= \frac{k_c q_i}{m_i} \sum_{j=1, j \neq i}^N \frac{q_j}{|\mathbf{r}_{ij}|^3} \mathbf{r}_{ij} + \mathbf{u}_{t_i} \end{aligned} \quad (6)$$

where k_c is the Coulomb constant, $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$ and \mathbf{u}_{e_i} is the resultant of the electrostatic forces acting on the i -th spacecraft. When the electrostatic force is exploited to provide part of the required control force, a coordination algorithm needs to be implemented because the dynamic of the swarm becomes highly coupled, i.e. a charge variation on the i -th spacecraft generates a variation in the forces acting on the other charged satellites. In this work we consider as main objective of the coordination scheme the reduction of total fuel consumption required to the whole swarm during the maneuver. Hence the following problem must be solved on line:

$$\forall t, \min_{\mathbf{q}} \sum_{i=1}^N |\mathbf{u}_{t_i}|. \quad (7)$$

Other figures of merit could be also considered such as balanced fuel consumption. A more in depth discussion of this second problem is given in [5].

3.1 The charge feedback

In this section we introduce a charge feedback law strategy that provides a solution to the problem in Eq.(7). On the basis of the results described in section 2, we assume that active emission of charged particles is used to produce fast variations of the spacecraft charge and

that a measurement of the level of charge of each satellite is available. Let us introduce the following charge feedback law for $i = 1, \dots, N$

$$I_i = \kappa_I [q_{des_i}(\mathbf{q}, \mathbf{r}_1, \dots, \mathbf{r}_N, \mathbf{u}_{des_i}) - q_i] \quad (8)$$

where κ_I is a scalar constant and q_{des_i} is the level of charge that the i -th spacecraft instantaneously wants to acquire and where $\mathbf{q} = [q_1, \dots, q_N]^T$ is the vector of the spacecraft actual charge values. There are different options to define the desired charge. In this work we assume that each spacecraft computes its individual desired charge trying to minimize its individual fuel expenditure e.g.

$$q_{des_i} = \mathbf{u}_{des_i} \cdot \mathbf{R}_i \mathbf{q} / |\mathbf{R}_i \mathbf{q}|^2 \quad (9)$$

where we define the actual configuration matrix \mathbf{R}_i as

$$\mathbf{R}_i = \frac{k_c}{m_i} \begin{bmatrix} \frac{\mathbf{r}_{i1}}{|\mathbf{r}_{i1}|^3}, & \dots, & 0, & \dots, & \frac{\mathbf{r}_{iN}}{|\mathbf{r}_{iN}|^3} \end{bmatrix}.$$

Note that, given the actual charges of the other swarm members in \mathbf{q} and the matrix \mathbf{R}_i , the direction of the electrostatic force acting on the i -th satellite is also given

$$\hat{\mathbf{u}}_{e_i} = \mathbf{R}_i \mathbf{q} / |\mathbf{R}_i \mathbf{q}|.$$

Then the charging strategy in Eq.(9) requires to assume that value of charge so that the resultant electrostatic force acting on the i -th spacecraft is the projection of \mathbf{u}_{des_i} onto $\hat{\mathbf{u}}_{e_i}$. A first attempt to study the stability properties of this feedback law and to relate the equilibrium position of the dynamical system in Eq.(8) to the solution of Eq.(7) is given in [1].

4 Simulation Results

In this section a simulation is proposed to illustrate the performances of the proposed charge feedback law. The simulation integrates the following dynamical system

$$\begin{aligned} \ddot{\mathbf{r}}_i + \mathbf{D}\dot{\mathbf{r}}_i + \mathbf{K}\mathbf{r}_i &= \mathbf{u}_{des_i} \\ \frac{dq_i}{dt} &= I_i = \kappa_I (q_{des_i} - q_i) \end{aligned} \quad (10)$$

with $i = 1, \dots, N$, and where, according to Eq.(6), we assume that the hybrid actuation system can instantaneously deliver the desired acceleration. The fuel savings triggered by the use of electrostatic actuation is clearly dependent from the maneuver that is performed i.e. from the time history of the desired acceleration \mathbf{u}_{des_i} generated by the path planning algorithm. In this paper

we described the derivation of the charging strategy assuming that a path planning algorithm that computes a suitable \mathbf{u}_{des_i} is already available. In particular, in the simulation below, the path planning technique introduced in [2] is used to calculate the desired acceleration vector for each spacecraft. It is shown in [5] how such path planning technique can be easily adapted to generate desired acceleration signals suitable for the application of the electrostatic actuation system. In Figure 4 the trajectories of a three spacecraft formation performing a deployment maneuver are shown. The satellites are supposed to have 50 kg mass and the center of the formation moves along a GEO orbit. Both the initial charges and positions have been randomly selected respectively in the sets $|q_i(t=0)| \leq 0.4 \mu\text{C}$ and $|r_i(t=0)| \leq 5 \text{ m}$ and are listed in Table 3. The final desired configuration

	Position(m)	Charge (μC)
SC1	[4.12, 0.84, -1.15]	-0.28
SC2	[-0.61, -1.65, -1.06]	-0.25
SC3	[-3.51, 0.81, 2.21]	0.2

TABLE 3. Initial conditions for the three spacecraft deployment maneuver.

is an equilateral triangle of approximately 35 m size. Finally, a saturation cycle has been included in the proposed charging strategy to limit the achievable maximum charge to $2 \mu\text{C}$.

In Figure 4 and in Figure 5 the three dimensional and planar view of the spacecraft trajectories during the maneuver are shown.

At the very beginning of the simulation the spacecraft are at rest condition and must acquire a certain radial velocity to reach the final desired position. Therefore in the first 150 seconds all the spacecraft assume a positive charge so that the electrostatic repulsion force can provide a considerable portion of the initial required acceleration. This is shown in the lower plot of Figure 6 where the time variations of the spacecraft charges at the very beginning of the simulation are displayed. After 250 s the satellites must start the deceleration process. For this reason attractive electrostatic forces must be established. The spacecraft marked with blue and green line experience during their motion a gravitational and inertial force along the x axis that, if not properly counteracted, would make their trajectories divergent. On the other hand the satellite marked with the red line

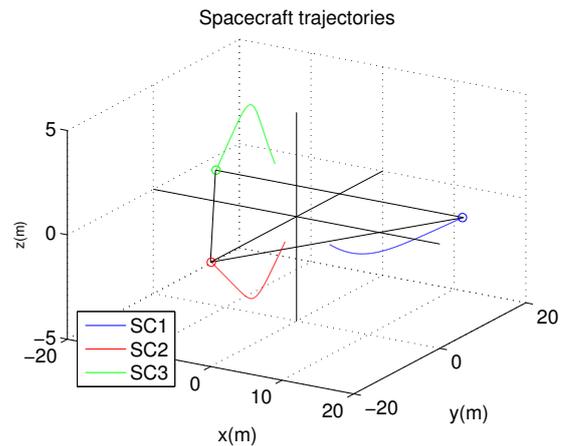


FIGURE 4. Three spacecraft deployment maneuver: 3D view.

moves approximately along the y axis where it experiences only an inertial force that pushes it toward the direction of the negative x. The charge time history for the whole simulation is presented in the upper plot of Figure 6. In order to better explain the behavior of the charging strategy, the electrostatic forces acting on the spacecraft after approximately 0.2 orbital periods of simulation are sketched as arrows in Figure 5 (electrostatic forces not in scale). This plot shows that the spacecraft marked in blue attracts the other two members of the swarm. This charge sign selection indeed allows to counteract part of the inertial and gravitational forces acting on the satellites with electrostatic forces. Moreover the charging strategy selects the charge magnitudes trying to minimize the residual thrusting acceleration needed to steer the swarm. The largest disturbing forces acting on the swarm are the gravitational and inertial forces along the x direction on the spacecraft marked in blue and green. In order to counteract this effect the blue and green spacecraft assume charges q_1 and q_3 respectively that yield a large charge product $q_1 q_3$. Therefore a large attractive electrostatic force is established between the two spacecraft. However, $q_1 > q_3$. Hence the attractive force between the spacecraft marked in blue and the one marked in red is larger in magnitude with respect to the repulsive force between the green and the red spacecraft. This implies that a portion of the electrostatic energy is also used to reduce the velocity of the spacecraft marked in red along the y direction. At the very end of the simulation all the spacecraft are again at rest condition. Note that the charge of the spacecraft marked in red converges to a null value since, at rest condition, no inertial force is

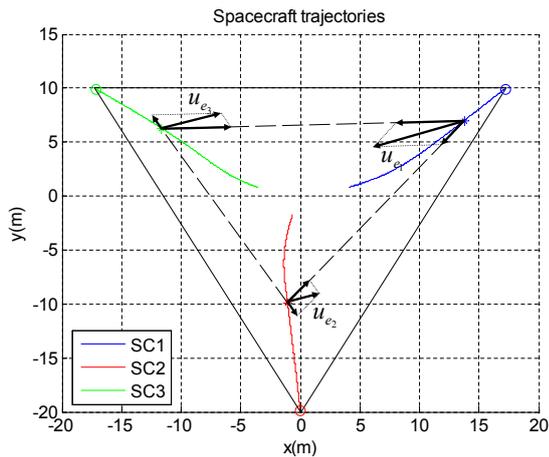


FIGURE 5. Three spacecraft deployment maneuver: planar view with frozen configuration after 0.2 orbital periods (electrostatic forces not in scale).

anymore acting on that spacecraft. In Figure 7 the Δv required to perform the whole maneuver with and without electrostatic actuation are displayed. The proposed charge control algorithm allows to save the 81% of the total Δv required to perform the maneuver.

5 Conclusions

In this work we investigate the possibility to steer a swarm of satellites by means of a hybrid thrusting and electrostatic actuation system. First a scheme for a possible charge actuation system has been developed. Simulations have shown that any desired level of charge in the given range can be fast achieved and maintained with high stability. A charge feedback law has been then introduced that, used in conjunction with a decentralized path planning algorithm, allows to increase the fuel efficiency of acquisition and reconfiguration maneuvers for swarm of satellites. In the future the following investigations are recommended for the further development of the concepts introduced in this work:

- Development of a dedicated charge emission device capable of emitting currents in large amperage and energy range.
- Development of a more detailed plasma-spacecraft interaction model in order to have a better estimate of the worst case scenarios especially for high spacecraft voltages.

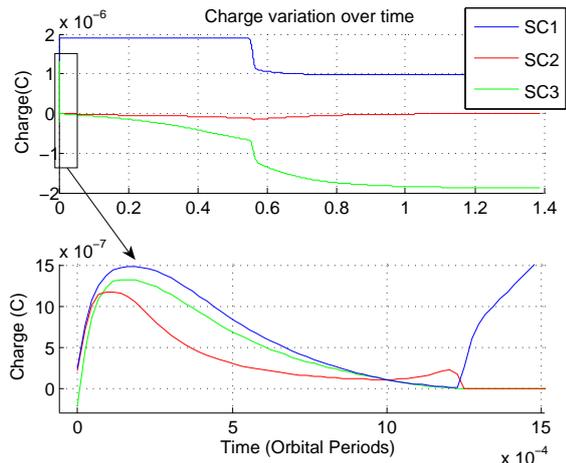


FIGURE 6. Three spacecraft deployment maneuver: charge variation.

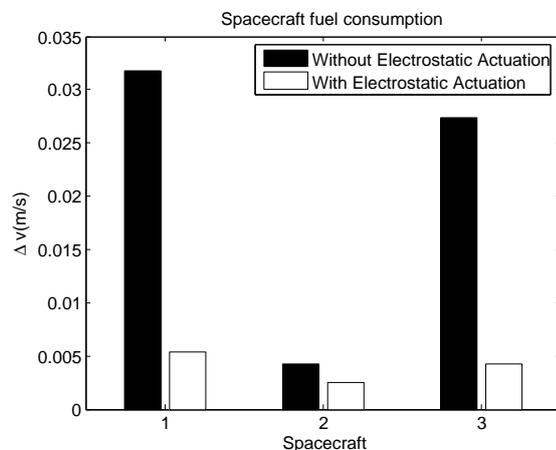


FIGURE 7. Fuel consumption of the three spacecraft deployment maneuver.

- Stability analysis of the charge feedback law in order to develop a more reliable algorithm for charge control.

References

- [1] D. Izzo and L. Pettazzi. Self-assembly of large structures in space using intersatellite Coulomb forces. In *57th International Astronautical Congress*, Valencia, Spain, 2006. IAC-06-C3.4/D3.4.07.
- [2] D. Izzo and L. Pettazzi. Autonomous and Dis-

- tributed Motion Planning for Satellite Swarm. *Journal of Guidance, Control and Dynamics*, 30(2):449–459, March–April 2007.
- [3] L. King, G. Parker, and J. Satwik, D. Chong. A study of inter-spacecraft coulomb forces and implications for formation flying. *Journal of Propulsion and Power*, 19(3):497–505, 2003.
- [4] G. Parker, L. King, and H. Schaub. Steered spacecraft deployment using interspacecraft coulomb forces. In *American Control Conference*, Minneapolis, Minnesota, June. 14–16 2006. Paper No. WeC10.5.
- [5] L. Pettazzi, H. Krüger, and S. Theil. Electrostatic forces for satellite swarm navigation and reconfiguration. Technical Report ARIADNA 05/4107a, September 2006. Available on-line [www.esa.int/gsp/ACT].
- [6] H. Schaub. Stabilization of Satellite Motion Relative to a Coulomb Spacecraft Formation. *Journal of Guidance, Control and Dynamics*, 28(6):1231–1239, November–December 2005.



Invariant Relative Satellite Motion

M. SABATINI^{*†}, D. IZZO[‡], G. PALMERINI[†][†]*Dip. Ingegneria Aerospaziale, Università di Roma La Sapienza*[‡]*Advanced Concepts Team, European Space Agency, Noordwijk, The Netherlands*

Abstract. Formation flying is a key technology enabling a number of missions which a single satellite cannot accomplish: from remote sensing to astronomy and fundamental physics. The design of relative navigation and control systems of the spacecraft is certainly one of the most challenging topics. In order to ease the tasks of these subsystems, a proper reference trajectory must be conceived, and a relative motion, which shows no drift even in presence of a large disturbance as the J_2 effect, could be a very attractive solution. This paper describes the research activities for finding invariant relative orbits under J_2 effect. Numerical tools as genetic algorithms enabled the discovery of two special inclinations which represent the necessary conditions for periodicity of the motion. These results generated interest, and analytical explanations for the numerical evidence have been proposed: on-going studies face this problem from different points of view, and basic results are reported.

1 Introduction

Formation flying is a key technology enabling a number of missions which a single satellite cannot accomplish: from remote sensing to astronomy and basic physics. In order to keep the satellites of the formation in the de-

signed configuration, and therefore to achieve the mission's goals, control actions are needed. The cost of this orbital control in terms of ΔV limits both the mission duration and the expected performances. Advantageous dynamics could reduce the cost of these operations, in particular the possibility to obtain periodic or quasi-periodic natural relative motion would be a significant saving factor. Many different approaches to find a periodical relative motion are considered in the recent literature. Inalhan, Tillerson and How [3] found the analytical expression for the initial conditions resulting in periodic motion based on the classical Tschauner-Hempel equations [9]. Kasdin and Kolesman [5] used the epicyclic orbital elements theory to derive bounded, periodic orbits in presence of various perturbations. Vaddi, Vadali, and Alfriend [12] studied the Hill-Clohessy-Wiltshire [2] (HCW) modified system to include second order terms. Finally, Schaub and Alfriend [8] formulated the conditions for invariant J_2 relative orbits introducing relations between the mean orbital elements of the two satellites. The analytical approaches taken in these works lead to initial conditions that ensure exact periodicity in approximated dynamical models or initial conditions resulting in bounded (but not periodic) relative motion in more detailed dynamical models. In the frame of the Ariadna project carried out by the Advanced Concepts Team of the European Space Agency and the Department of Aerospace Engineering of the University of Rome [6], the possibility to

^{*}Corresponding author. E-mail address: marcosabatini@hotmail.it

obtain natural periodic relative motion of formation flying LEO (Low Earth Orbits) satellites has been investigated numerically. The algorithm is based on a genetic strategy (GA), refined by means of nonlinear programming, that rewards periodic relative trajectories. Only the J_2 perturbed case is considered, as presence of a dissipative disturbance like drag does not enable a periodic motion. Using the proposed numerical approach, it has been possible to find two couples of inclinations (63.4 and 116.6, the critical inclinations, and 49 and 131, two new "special" inclinations) that seemed to be favoured by the dynamical system for obtaining periodic relative motion at small eccentricities [7].

This interesting numerical result still missed a mathematical or physical explanation. Vadali, Sengupta, Yan, and Alfriend in [10] later found that these special inclinations are also the values for inclination which enable equal in-plane and out-of-plane fundamental frequencies, resulting in non-precessing and distortion-free relative orbits over the short-run.

Following these results, a new kind of study has been started, on mathematical basis, which aims to show that:

1. a truly periodic relative motion in a J_2 perturbed environment is generally not possible
2. two special inclinations exist for which at least a four dimensional subspace of the state space shows superior quasi-periodicity features

2 The numerical approach: definition of the problem

Consider a generic non-autonomous dynamical system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t)$, e.g. the relative dynamics of satellites flying in formation. Define $\delta\mathbf{x} = \mathbf{x}_0 - \mathbf{x}(T)$, where \mathbf{x}_0 is the system state at the initial time and T is a time variable here called 'candidate period' for reasons that will soon be clear. Then, the following optimisation problem is defined:

$$\begin{cases} \text{find: } & \boldsymbol{\kappa} = [\mathbf{x}_0, T]^T \\ \text{to maximise: } & J = J(|\delta\mathbf{x}|) \\ \text{subject to: } & \dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t) \end{cases} \quad (1)$$

where the objective function J is constructed in such a way as to have its global maximum at $\delta\mathbf{x} = 0$. The optimisation problem above is equivalent to the task of finding as-periodic-as-possible solutions to the system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t)$. These solutions correspond, in our case, to minimal relative orbit drift. As we study a number of

Number of individuals	20
Number of generations	500
Number of significant digits	9
Crossover Probability	0.85
Initial Mutation Rate	0.005
Minimum Mutation Rate	0.0005
Maximum Mutation Rate	1
Reproduction Plan	Steady-State Replace-Worst

TABLE I. Parameters used in PIKALIA.

systems $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t)$, we face different levels of complexity for the optimisation and for objective function properties. Think about the relative motion between satellites moving on keplerian orbits, the problem defined by Eq. (1) has an infinite number of solutions, corresponding to orbits with equal semi-major axis. A similar structure is also expected when the keplerian dynamic is perturbed. As a consequence, a genetic approach, avoiding issues related to domain knowledge and able to cope with multiple local and global minima, has been selected to perform a search in the solution space. The PIKALIA freely available software [1] was used in this work as genetic optimiser. Table 1 shows the fundamental parameters of the genetic algorithm used in all the simulations.

The best solution returned by the genetic algorithm is then refined locally by means of a nonlinear programming solver. In our simulations the decision vector \mathbf{k} contains the initial relative position, the initial relative velocity, and the candidate period T . We consider the relative motion between two satellites: a chief and a deputy to use a popular terminology connected to formation flying research. The absolute dynamics of both the chief and the deputy are simulated propagating the inertial coordinates of the spacecraft in time, $\frac{d^2\mathbf{r}}{dt^2} = -\frac{\mu}{r^3}\mathbf{r} + \mathbf{P}$ where \mathbf{P} are the perturbing action considered, μ is the planetary gravitational parameter and \mathbf{r} is the position vector. The relative state is then evaluated by means of Eq. (2):

$$\begin{cases} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \mathbf{R} \left(\begin{bmatrix} X_d \\ Y_d \\ Z_d \end{bmatrix} - \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} \right) \\ \begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix} = \mathbf{R} \left(\begin{bmatrix} \dot{X}_d \\ \dot{Y}_d \\ \dot{Z}_d \end{bmatrix} - \begin{bmatrix} \dot{X}_c \\ \dot{Y}_c \\ \dot{Z}_c \end{bmatrix} \right) - \boldsymbol{\omega} \wedge \begin{bmatrix} x \\ y \\ z \end{bmatrix} \end{cases} \quad (2)$$

where \mathbf{R} is the rotation matrix from the inertial coordinate system to the Local Vertical Local Horizontal (LVLH) frame in which the relative state is defined. The subscripts c and d stand for chief and deputy satellites. The orbit of the chief is considered known and the initial conditions to propagate the deputy motion are obtained transforming the relative $\delta\mathbf{x}_0$ position into absolute coordinates inverting Eq. (2).

We then defined $T = T_{kep} \pm \bar{\kappa}_7 k$, where k is a properly chosen constant (some tens of seconds) and T_{kep} is the orbital period $2\pi\sqrt{a^3/\mu}$ of the chief orbit. T is clearly a crucial parameter. At $t = T$, the final relative coordinates are compared to the initial relative coordinates, thus determining the quality of the individual. A good individual has a small $\delta\mathbf{x}$ and its position in the individual ranking is high, therefore it has a larger chance to mate and to generate "good" sons. Its genes will survive in the next generation, and if they will be ranked first in the last generation, they will be further refined by a local optimiser and represent the set of initial conditions that generate the minimum drift relative orbit. The fitness function we used to rank the individuals is:

$$J(\boldsymbol{\kappa}) = \frac{1}{0.001 + \sum_i \frac{\delta x_i}{x_0}} \quad (3)$$

A perfect individual (periodic motion) has a fitness value of 1000, while a percentage difference of 0.1% between the initial and final state, brings down the fitness value to 500, a difference of 1% corresponds to a fitness value of 90, and so on.

3 The numerical approach: results

Let us study the solutions of Eq. (1) in the case \mathbf{f} describes the relative motion between two satellites orbiting around an oblate Earth. As already mentioned, this corresponds to minimising the relative orbital drift. Some previous work has been done to determine the possibility of invariant relative satellite motion when J_2 is considered as a perturbing term. In particular, the paper by Schaub and Alfriend [8] introduces the so called

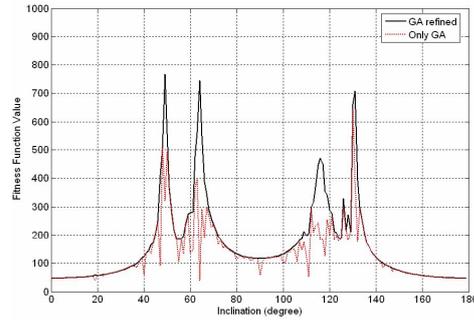


FIGURE 1. Best Individual Fitness Value vs. Inclination of Reference Orbit.

J_2 invariant relative orbits. In their work, mean orbital elements are used and the secular drifts of the longitude of the ascending node and of the sum of the argument of perigee and mean anomaly are set to be equal between two neighbouring orbits. Even though called J_2 invariant orbits, these two conditions are only valid as a first order approximation.

We use our numerical approach based on the solution of the optimisation problem stated in Eq. (1) to check to what extent the residual drift obtained with this analytical approach is an artifact of the use of mean elements. Repeating the calculation for the entire range of inclinations, the results vary sensibly, disclosing a previously unknown feature of the invariant relative motion. In Figure 1 we report the best fitness function reached for different inclinations ranging from 0 to 180 degrees. The other orbital parameters of the Chief satellite used for this simulation were $a = 6678$ km, $e = 0.00118$, $\omega = 90$, $\Omega = 270$, $\theta = 0$, having used a widely common symbolism. In Figure 1 we report both the output from the genetic algorithm and the final solution obtained refining the solution with a local optimisation.

For all inclination the minimal drift is not zero, with two remarkable exceptions: 49 and 63.4, and their symmetric counterparts (with respect to 90), i.e., 131 and 116.6. In the following we will refer to these as "special inclinations". The heuristic of the genetic algorithms is definitely not responsible for these peaks, as it turns out by actually propagating the resulting best individuals. At a generic inclination, say 35, the best individual returned by the optimisation results in a relative motion that is clearly not periodic, as visualised in Figure 3. The small residual drift is comparable to the one that results using Schaub J_2 -invariant orbit condition.

At the special inclinations 63.4 and 116.6 the rela-

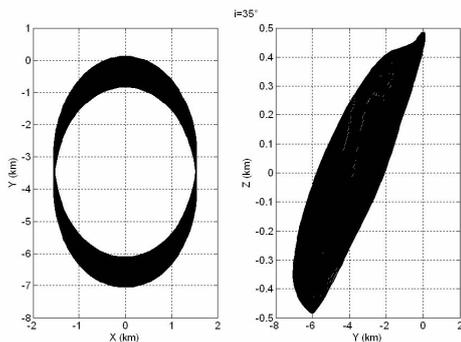


FIGURE 2. 100 Relative Orbits for a Perturbed case at 35 Inclination (Best Individual).

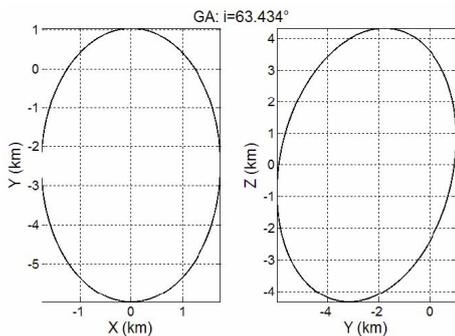


FIGURE 3. 100 Relative Orbits for a Perturbed case at 63 Inclination (Best Individual).

tive motion turns out to have features typical of pure periodic motion (see Figure 3). These features are probably related to the cancellation of the secular drift of the perigee argument, which causes the variation of all parameters to happen with the same main frequency. At the other two special inclinations (49, 131) the relative satellite motion resulting from the best individual has only a very small drift, as shown in Figure 4.

The small residual drift does not allow us to conclude that the motion is perfectly periodical at these inclinations. We were in fact unable to find a fitness value of 1000 (meaning perfect periodicity, according to Eq. (3)) at any inclination. And we will indeed see that the motion (at least in the linear system approximation) is not periodical at any inclination, if we take the mathematical definition of periodicity.

A more detailed plot of the objective function achieved around the special inclination is shown in Figure 5. The clutter that can be observed in the graph is a

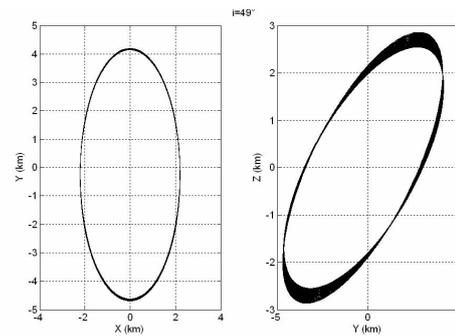


FIGURE 4. Relative Orbits for a Perturbed case at 49 Inclination (Best Individual).

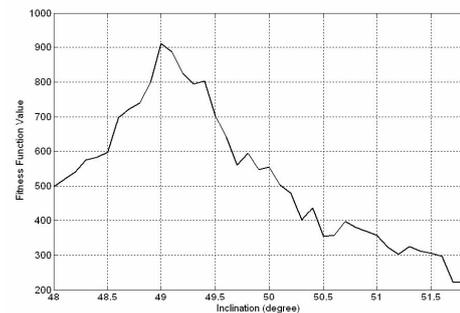


FIGURE 5. Details around a special inclination.

consequence of the numerical optimisation process, amplified by the definition of the objective function given by Eq. (3). At higher values of the fitness very small differences in the residual drift cause significant differences in the objective function value.

For completeness we also report in Figure 6 a plot of the period of the found minimum drift orbits. This is clearly quite different from the keplerian period confirming the importance of having let the optimiser to choose it.

The existence of the two special inclinations certainly represented the main finding of the Ariadna project carried out by the Advanced Concepts Team of the European Space Agency and the Department of Aerospace Engineering of the University of Rome [6], but it clearly called for some physical or analytical explanation.

4 An analytical explanation

Vadali et al. in [10] arrived at a physical explanation to the special inclinations. The first step consists in a

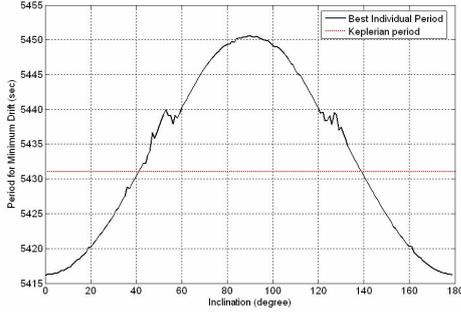


FIGURE 6. Details around a special inclination.

geometric description of the relative motion using the unit-sphere approach. The expressions for the in-plane and cross-track motion variables are linearized to extract the fundamental frequencies of motion. A set of classical differential orbital element initial condition formulae are derived, valid even for circular orbits. In this way the difference between the in-plane and cross-track frequencies (n_{xy} and n_z , respectively), over a short time interval (compared to the period of the differential nodal precession rate), is evaluated as:

$$n_z - n_{xy} = \dot{\omega}_1 - \frac{\sin i_0 \Delta \dot{\Omega} \Delta i}{\Delta i^2 (\sin i_0 \Delta \Omega(0))^2} \quad (4)$$

where i_0 is the reference orbit inclination, $\Delta \dot{\Omega} \Delta \Omega \Delta i$ the differences in RAAN rate, RAAN and inclination between the formation members. The drift rate for the argument of perigee can be written as

$$\dot{\omega}_0 = -k \left(2 - \frac{5}{2} \sin^2 i_0 \right) n_0 \quad (5)$$

where:

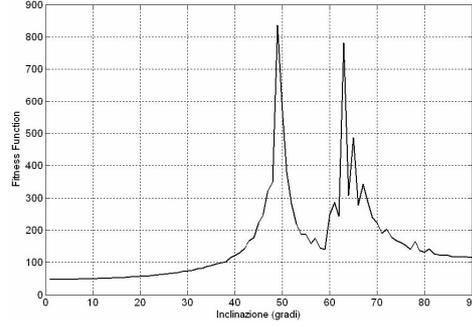
$$k = -\frac{1}{5} J_2 \left(\frac{R_E}{a_0} \right)^2 n_0 \quad (6)$$

The differential nodal precession rate for near-circular orbits is

$$\Delta \dot{\Omega} = -k \sin i_0 \Delta i \quad (7)$$

Neglecting the effect of $\Delta \omega$, the frequency mismatch is estimated as follows:

$$n_z - n_{xy} = k \sin^2 i_0 \left(\frac{5}{2} + \frac{\Delta i^2}{\Delta i^2 + (\sin i_0 \Delta \Omega(0))^2} \right) \quad (8)$$


 FIGURE 7. Results of the application of the numerical method applied to the linear J_2 model.

For the special case of the projected circular orbit (ref. [11]), it can be shown that:

$$\begin{aligned} \Delta \Omega(0) &= -\frac{\rho(0)}{\alpha_0 \sin i_0} \sin \alpha(0) \\ \Delta i &= \frac{\rho(0) \cos \alpha(0)}{\alpha_0} \end{aligned} \quad (9)$$

where $\alpha(0)$, is the desired initial phase angle and $\rho(0)$ is the initial radius of the relative orbit in the y - z plane. By substituting Eqs. (9) into Eq. (8), the frequency matching condition is satisfied by two possible values of the inclination:

$$i^* = \sin^{-1} \left(\sqrt{\frac{2}{2.5 + \cos^2 \alpha(0)}} \right) \quad (10)$$

The inclinations for $\alpha(0) = 0$ are $i^* = 49.11$ and $i^* = 130.89$ and for $\alpha(0) = 90$, the results are the critical inclination values: $i^* = 63.43$ and $i^* = 116.57$. Anyway, $\alpha(0)$ is a parameter varying between 0 and 360 degrees, this means that $i^* \in [49.11, 63.43]$, while the numerical evidence is for i^* equal to the boundary values of the interval.

5 An alternative analytical approach

In [4], a linear model is obtained, describing the relative dynamics of a formation in a J_2 perturbed environment (circular reference orbit case). If the numerical optimization performed for the nonlinear formation dynamics is used now for the linear model, similar results are obtained, as Figure 7 confirms.

It is therefore possible to focus the attention on the linearized dynamics, since the causes of the existence of special inclinations are in the first order gradients of gravity and J_2 disturbance as shown by the results obtained. The adoption of a linear system as object of the

study allows the use of relevant theorems valid for linear systems. Therefore, the time behaviour of the six-eigen values λ of $\Phi(t)$ becomes quite crucial. $\Phi(t)$ is evaluated by means of:

$$\begin{aligned} \frac{d\Phi(t)}{dt} &= \mathbf{A}(t)\Phi(t) \\ \Phi(0) &= \mathbf{I} \end{aligned} \quad (11)$$

The $\mathbf{A}(t)$ state matrix is a 6-by-6 matrix: a relative periodic motion is possible only if all the eigen values of $\Phi(t^*)$ are equal to 1 at $t = T$ (i.e. the imaginary and real parts are both zero). This is never the case at any inclination in the linear system adopted. Nevertheless there are some inclinations where, at time instants close to the orbital period T , four eigen values are equal to 1. As an example see Figure 8, referred to the $i=45^\circ$ case. The imaginary part of the eigen values is never 0 for more than a pair of eigen values at the same time, therefore with multiplicity greater than two is never achieved.

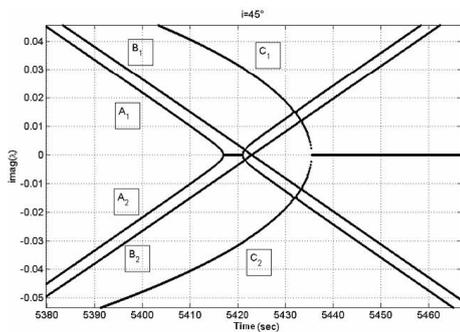


FIGURE 8. Time behaviour of the imaginary part of the 6 eigen values of $\Phi(t)$: 45 degrees case.

However the pairs A_1, A_2 and B_1, B_2 have an imaginary part which is zero for time instants which get closer as the inclination increases and eventually coinciding at the inclination value equal exactly to the special inclination 49.1° (Figure 9). Figure 10 confirms that not only 4 eigen values have a zero imaginary part in a certain instant, but that in that instant also the real part of the same eigen values is equal to 1.

Similar considerations can be done for the other special inclination, 63.1 : at these two special inclinations the motion of the formation is by far more stable than at any other inclination, though it is not mathematically periodic in its six components and just in four are close to be periodic sharing the same repetition period. In fact, for inclinations in the range $[49.1, 63.1]$, while 4 eigenvalues with imaginary part equal to zero are possi-

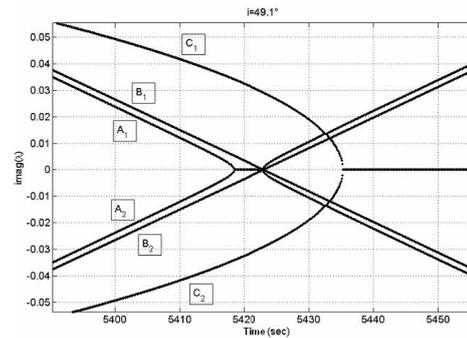


FIGURE 9. Time behaviour of the imaginary part of the 6 eigen values of $\Phi(t)$: 49.11 degrees case.

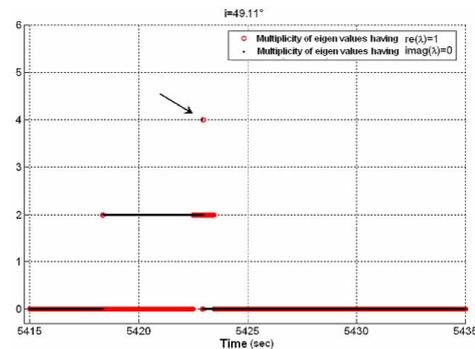


FIGURE 10. Multiplicity of the eigen values with zero imaginary part and that of eigenvalues with real part equal to 1: 49.11 case

ble, they do not have real part equal to one (Figure 11 and Figure 12). For inclinations greater than 63.1° , even if it is possible that 4 eigen values have imaginary part equal to zero, again they do not have real parts equal to one. These analytical observation allow us to conclude that the key to explain the special inclination phenomena is in the spectral properties of the matrix $\Phi(t)$.

6 Future Developments

A quasi-periodic relative orbit represents a possibly very low-cost station keeping solution. Even though it is true that J_2 is not the only environmental disturbance, it is certainly the largest for a wide range of missions profiles. In addition, the special inclination 49.1 is quite promising since it is not very different from the inclination of GPS, of the future GALILEO, of the International Space Station. Having an almost free-of-control formation could suggest a number of solutions, different

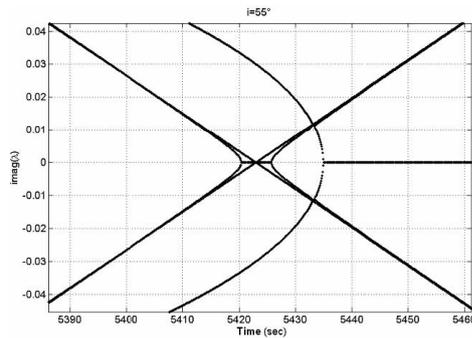


FIGURE 11. Time behaviour of the imaginary part of the 6 eigenvalues of $\Phi(t)$: 55 degrees case.

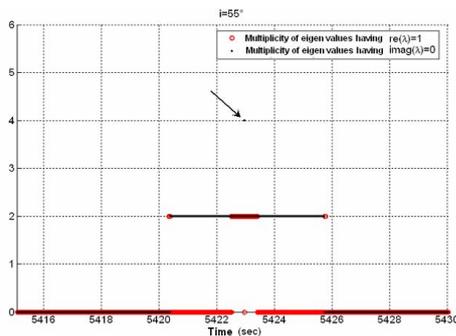


FIGURE 12. Multiplicity of the eigen values with zero imaginary part and that of eigenvalues with real part equal to 1: 55 case.

from the usual formation tasks: it can be useful for occupying the same orbital slot without conflicting, serving as a spare satellite, or for continuous monitoring of the health of the main satellite. Furthermore, once invariant J_2 orbits have been assessed, the effects of other perturbations can be isolated and contrasted in a most efficient way, if needed, or alternatively, they can be fully exploited on purpose. The use of the special inclinations for future formation flying missions, thus rest on further understanding the benefits that these particular orbits have to offer in terms of ΔV savings.

7 Conclusions

In this brief communication we have shown that the dynamic resonance phenomenon observed numerically during the Ariadna project carried out by the Advanced Concepts Team of the European Space Agency and the Department of Aerospace Engineering of the University of Rome [6] can be explained by looking at the spectral

properties of the dynamic matrix in a carefully linearised set of equations. No periodic motion is possible at any inclination value, yet at the special inclination values and only at those values, a four dimensional subspace of state space is very close to be periodical.

References

- [1] P. Charbonneau and B. Knapp. A User's Guide to PIKAIA 1.0. *NCAR Technical Note*, 418, 1995.
- [2] W. Clohessy and R. Wiltshire. Terminal Guidance System for Satellite Rendezvous. *Journal of the Aerospace Sciences*, 27(9):653–658, 1960.
- [3] G. Inalhan, M. Tillerson, and J. How. Relative dynamics and control of spacecraft formations in eccentric orbits. *Journal of Guidance, Control, and Dynamics*, 25(1):48–59, 2002.
- [4] D. Izzo, M. Sabatini, and C. Valente. A new linear model describing formation flying dynamics under J_2 effects. In *Proceedings of 17th AIDAA national congress held in Rome, ITA*, pages 15–19, 2003.
- [5] N. Kasdin and E. Kolemen. Bounded, Periodic Relative Motion Using Canonical Epicyclic Orbital Elements (AAS 05-186). *Advances in the Astronautical Sciences*, 120(2):1381, 2005.
- [6] M. Sabatini, R. Bevilacqua, and M. Pantaleoni. A search for invariant relative satellite motion. Technical Report 04-4104a, European Space Agency, the Advanced Concepts Team, 2005. Available on line at www.esa.int/act.
- [7] M. Sabatini, D. Izzo, and R. Bevilacqua. Special Inclinations Allowing Minimal Drift Orbits for Formation Flying Satellites. *Journal of Guidance, Control, and Dynamics*, 31(1):94, 2008.
- [8] H. Schaub and K. Alfriend. J_2 Invariant Relative Orbits for Spacecraft Formations. *Celestial Mechanics and Dynamical Astronomy*, 79(2):77–95, 2001.
- [9] J. Tschauer and P. Hempel. Rendezvous zu einem in elliptischer Bahn umlaufenden Ziel. *Astronautica Acta*, 11(2):104–109, 1965.
- [10] S. Vadali, P. Sengupta, H. Yan, and K. Alfriend. On the Fundamental Frequencies of Satellite Relative Motion and Control of Formations (AAS

- 07-427). *Advances in the Astronautical Sciences*, 129(3):2747, 2008.
- [11] S. Vadali, S. Vaddi, and K. Alfriend. An intelligent control concept for formation flying satellites. *International Journal of Robust and Nonlinear Control*, 12(2-3):97–115, 2002.
- [12] S. Vaddi, S. Vadali, and K. Alfriend. Formation Flying: Accommodating Nonlinearity and Eccentricity Perturbations. *Journal of Guidance, Control, and Dynamics*, 26(2):214–223, 2003.



Pulsar Navigation

JOSEP SALA^{*†}, ANDREU URRUELA[†], XAVIER VILLARES[†], JORDI ROMEU[†],
SEBASTIÀ BLANCH[†], ROBERT ESTALELLA[‡], JOSEP M. PAREDES[‡]

[†]*Dpt. of Signal Theory and Communications, Technical University of Catalonia (UPC),*

c/Jordi Girona 1-3, Campus Nord UPC, D5-114b, 08034 Barcelona, Spain

[‡]*Dpt. of Astronomy and Meteorology, University of Barcelona (UB),*

Diagonal 647, E-08028 Barcelona, Spain

Abstract. This paper evaluates the fundamental performance bounds, signal processing and technological complexity associated with the synchronization to radio and X-ray pulsars as well as its impact on the positioning accuracy of an autonomous spacecraft navigation system. Performance analysis of synchronization and location algorithms builds on Maximum Likelihood (ML) estimation which provides, asymptotically, unbiased minimum variance estimates with mean square error strictly approaching the Cramer-Rao Lower Bound (CRLB).

1 Introduction

Pulsar grids have been suggested in the past for autonomous, earth-independent, deep space location and for the establishment of a universal time base. This paper summarizes a feasibility study of deep space navigation based on radio and X-ray pulsar timing information. The analysis is focused on the signal processing aspects of the system according to the structure in figure (1). Therefore, the quality of measurements attainable with on-board instrumentation is evaluated for realistic space scenarios for the selected pulsars.

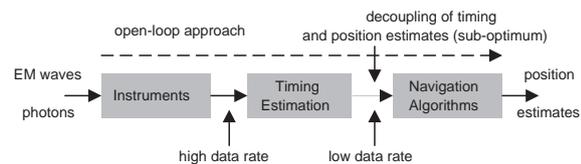


FIGURE 1. Navigation system structure. The first stage (instruments) converts EM waves/photons to raw digital input signals. Decoupling of timing (second stage) and location (third stage) parameter estimation is considered for ease of implementation and obtention of closed-form expressions of performance measures. Timing estimation uses information from a pulsar database to compare measured with expected pulse arrival times. The location stage uses this differential information to derive a position estimate.

Radio pulsars [5] are broadband stellar pulsating radio sources powered by the rotation of a neutron star, implying a great stability of the pulsar period (in the millisecond to second range), which steadily increases as the pulsar releases its rotational energy. Some 1300 pulsars are known [6], and more are discovered in every new survey [3]. Most of millisecond pulsars belong to binary systems, their pulse arrival time affected by the system orbital motion. In some pulsars, irregular-

^{*}Corresponding author. E-mail address: josep.sala@upc.edu

ities (glitches) in their rotation speed ω have been observed every few years, with fractional changes in the order of 10^{-6} for the Vela pulsar, while only 10^{-8} for the Crab pulsar. Pulsars emit a pulse of radiation during a small fraction of the rotation period. Although individual pulses vary in strength and shape (micro-structure), the average pulse shape is stable and characteristic of each pulsar. Very precise models have been established for the mean arrival time of pulsars whose extraordinary stability outperforms the most precise artificial time-bases. The most stable are millisecond pulsars. Unfortunately, their SNR is considerably poorer than for other longer period pulsars. Radio pulsars have steep power-law spectra, with spectral indices α ($S_\nu \propto \nu^\alpha$) ranging from -3.0 to -0.1 , with a median of ~ -1.7 . The strongest pulsar at 1 GHz has a flux density of 1.7 Jy ($1 \text{ Jansky} = 10^{-26} \text{ W m}^{-2} \text{ Hz}^{-1}$).

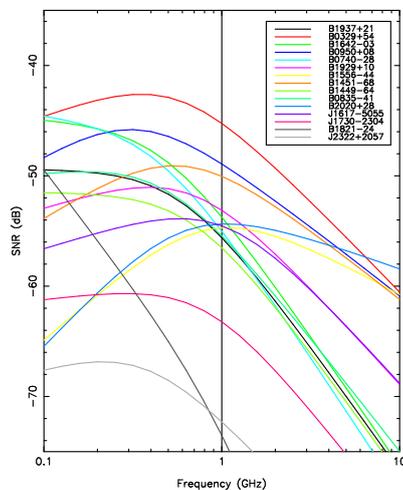


FIGURE 2. SNR as a function of frequency for the observation of the 15 best Q pulsars, assuming an effective area $A_e = 10 \text{ m}^2$, a beam efficiency $\eta_M = 0.9$, and no contribution from the Sun emission.

X-ray pulsars can be grouped in two different families according to the powering source: accretion- and rotation-powered pulsars. Accretion-powered pulsars are found in X-ray binary systems, the brightest class of X-ray sources in the sky. An X-ray binary contains either a neutron star or a black hole accreting material from a companion star. A neutron star with a strong magnetic field ($\sim 10^{12} \text{ G}$) will disrupt the accretion flow at several hundred neutron star radii and funnel material

onto the magnetic poles. If the magnetic and rotation axes are misaligned, X-ray pulsations will be observed if the beamed emission from the magnetic poles rotates through the line of sight. Thirty-two accretion-powered X-ray pulsars have been discovered [8], with pulse periods distributed between 0.069 s and 835 s. Long term monitoring has revealed different types of behaviour for the pulse period: a) linear decrease with time (spin-up) with erratic variations around the trend, b) no long term trend and only a random walk in the period, c) a steady increase in the period (spin-down). The pulse period fluctuations reflect inhomogeneities in the accretion flow, which excludes these pulsars as a stable timing source. Rotation-powered pulsars are rapidly spinning, strongly magnetized neutron stars radiating at the expense of rotational energy. For a long time the most luminous of all rotation-powered pulsars, the Crab pulsar, had been the only radio pulsar detected at X-ray energies. Today, as a result of observations with ROSAT and ASCA, 27 pulsars have been detected [1], with a range of spin periods between 1.6 ms and 530 ms. However, some of them have shown lack of pulsed emission at X-rays and/or the presence of glitches. This excluded them from our list of suitable candidates. Ten X-ray pulsars were finally selected [7]. Millisecond pulsars are distinguished by their small spin periods ($P \leq 20 \text{ ms}$) and high rotational stability ($dP/dt \simeq 10^{-18} - 10^{-21}$). Among the selected rotation-powered pulsars with pulsed X-ray emission, there are four millisecond pulsars (PSR J0437-47, PSR B1821-24, PSR J2124-33 and PSR J0030+0451). PSR J0030+0451 was recently discovered, presenting a period of 4.8 ms and gross similarity between the radio and the X-ray pulse profile (Becker et al. 2001). In the case of PSR B1821-24, the pulse profile has narrow peaks while for PSR J0437-47 and PSR J2124-33 the pulse profiles are broad.

2 Current Research

In this section we describe accomplished research on: (a) the fundamentals of radio- and X-ray pulsar synchronization, which, according to the structure in figure (1), is the previous step to navigation algorithms; (b) a simple navigation algorithm suitable for feasibility analysis, which incorporates ambiguity resolution. Conclusions and future vision will be provided in forthcoming sections.

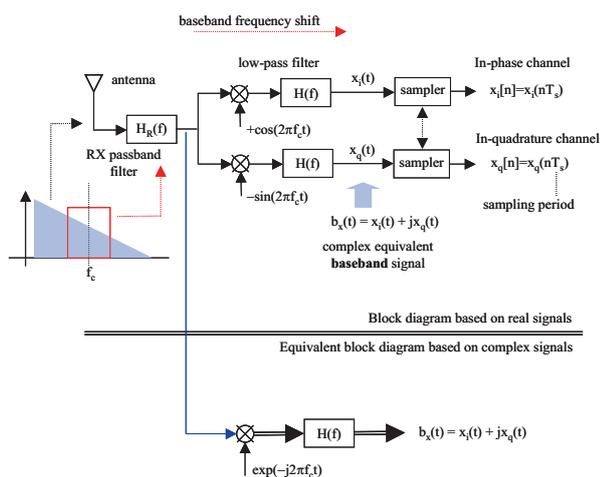


FIGURE 3. (Above) Signal processing scheme (IQ-sampling) for generating the in-phase and in-quadrature components of a pass-band radio signal at the central frequency f_c . (Below) Equivalent representation in complex notation.

2.1 Radio Pulsar Synchronization

Due to pulsar rotation, the noiseless pulsar signal $x_p(t)$ is described as a random signal with periodically time-varying statistics. That is, $x_p(t)$ is a cyclo-stationary stochastic process [2], with auto-correlation function $r_{x_p x_p}(t, \tau) = \mathbb{E}[x_p^*(t)x_p(t + \tau)]$ depending on the delay variable τ and periodic in the time variable t , with $\mathbb{E}[\cdot]$ the expectation operator. We denote the inverse of this rotation frequency the Pulse Repetition Period (PRP) T . Then, the k -th period of $x_p(t)$, corresponding to the time interval $[-T/2 + kT, T/2 + kT)$, displays a broadband micro-structure constituted by a number of random sub-pulses. The instantaneous power profile of $x_p(t)$ is defined as,

$$\sigma_p^2(t) = r_{x_p x_p}(t; 0) = \mathbb{E}|x_p(t)|^2 \quad (1)$$

After propagating through the interstellar medium, $x_p(t)$ experiences a frequency dependent time delay $T(f) = T_\infty - D(r)f^{-2}$ that distorts the broadband pulse envelope, with the dispersion parameter $D(r)$ depending on the distance r from the pulsar. As $x_p(t)$ extends over many octaves, multi-channel processing for pulsar observation is usually carried out over a fraction of its broadband spectrum. Considering a single channel, a portion of that spectrum is pass-band filtered at the central frequency of the antenna: $x_h(t) = h_R(t) * x_p(t)$ (with $*$ the convolution operator and $h_R(t)$ the front-end filter impulse response), converted to base-band

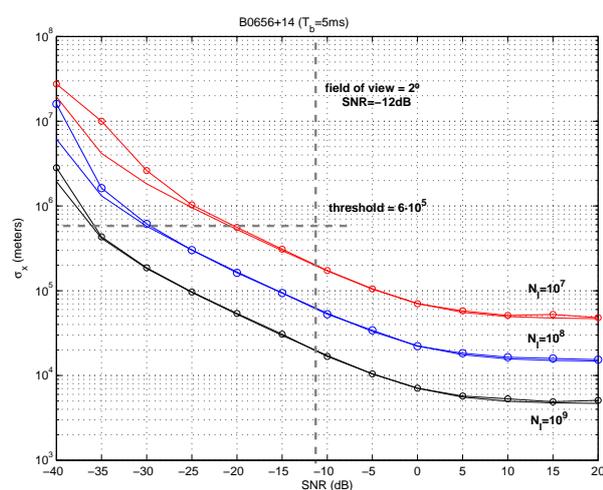


FIGURE 4. Standard deviation of the range estimator (in meters) for the B0656+14 X-Ray pulsar as a function of the SNR. Different integration times are considered and both the ML estimator (round markers) and the related CRB bound (no markers) are depicted. The low-SNR threshold is also shown in the figure.

from the channel central frequency f_c and low-pass filtered by the analysis filter impulse response $h_b(t)$,

$$b(t) = h_b(t) * (e^{-j2\pi f_c t} \cdot x_h(t)) \quad (2)$$

to deliver the complex equivalent baseband signal $b(t)$, as shown in figure (3). Consecutive samples of $b(t)$, $b[n] = b(nT_s)$, with T_s the sampling period, are stored in vector \mathbf{x} . In fact, the pulsar signal $x_h(t)$ is immersed in radio noise from the Sun and the Galaxy, so that incorporating the equivalent baseband noise into \mathbf{x} , we get $\mathbf{x} = \mathbf{p}_\theta + \mathbf{n}$, with \mathbf{p}_θ and \mathbf{n} vectors of samples of the pulsar and noise processes, respectively; $\theta = [\tau]$ constitutes the vector of unknown pulsar parameters, with τ the arrival phase of the pulsar signal with respect to the local timing reference. Although a more elaborate model might consider $\theta = [\tau; \delta_\tau]$, where the extra parameter δ_τ models pulsar period errors (frequency errors) in the local timing reference, meaningful feasibility results can already be obtained from this simplified model.

Pulsar phase synchronization constitutes the first stage in the navigation algorithms. In establishing the performance of pulsar synchronization, we assume the low-SNR Maximum Likelihood (ML) Estimator under complete knowledge of the pulsar and noise statistics, which is, asymptotically, a minimum variance unbiased

ased estimator [4]. This estimator maximizes the probability density of the received signal conditioned on the unknown parameters (pulsar arrival phase): $p(\mathbf{x}|\tau)$. At low SNR (small antenna), its performance is controlled solely by the second order statistics (correlation matrix) of the pulsar signal, which can be (partially) inferred from experimental results. Hence, the pulsar instantaneous power profile $\sigma_p^2(t)$ is used as the only available information. Therefore, under Gaussian noise, the final phase estimate is found to be,

$$\hat{\tau} = \underset{\tau}{\operatorname{argmax}} (\mathbf{x}^H \mathbf{R}_n^{-1} \mathbf{R}_p(\tau) \mathbf{R}_n^{-1} \mathbf{x}) \quad (3)$$

where \mathbf{x} is a vector containing contiguous samples of the baseband received signal, \mathbf{R}_n is the correlation matrix of noise and $\mathbf{R}_p(\tau)$ is a model for the correlation matrix of the pulsar signal depending on the unknown phase parameter τ , which contains $\sigma_p^2(t)$ in its diagonal. This equation is applicable whatever the structure of \mathbf{R}_n and $\mathbf{R}_p(\tau)$, which makes it suitable for the multiband case. The phase estimator becomes the maximization of $J(\tau)$, where,

$$J(\tau) = \sum_{i=-N_1}^{+N_1} \operatorname{SNR}_i \cdot \sum_n |x_i[n]|^2 \sigma_b^2(\tau - \phi_i + nT_s)$$

with SNR_i the signal to noise ratio in the i -th subband, $N_{sb} = 2N_1 + 1$ the number of subbands, $x_i[n]$ the sampled signal in each subband and ϕ_i the pulsar phase delay due to interstellar dispersion in each subband.

The Cramer-Rao Lower Bound (CRLB) is a lower bound to the variance of the unbiased estimate of a parameter [4]. As the size of the data record grows to infinity, the CRLB constitutes a strict performance measure of the (asymptotically unbiased) ML estimator. For an unbiased estimate $\hat{\tau}_x$ of τ , we have,

$$\operatorname{CRLB}(\tau) = \mathbb{E}_{\mathbf{x}|\hat{\tau}_x} |\hat{\tau}_x - \tau|^2 \geq \frac{-1}{\mathbb{E}_{\mathbf{x}} \nabla_{\tau}^2 \ln p(\mathbf{x}|\tau)}$$

in terms of the conditional data probability density $p(\mathbf{x}|\tau)$ evaluated at the true parameter τ . For a multiband processor the CRLB of the pulsar phase can be found expressed by the following equality,

$$\operatorname{CRLB}_{mb}(\tau) = \frac{1}{4\pi^2 T_o B} \left(\frac{\sigma_n^2}{\sigma_{p,\tau}^2} \right)^2 \cdot \gamma_p^2 \cdot \xi_p^2 \quad (4)$$

with T_o the observation time, B the global bandwidth, σ_n^2 the noise power, $\sigma_{p,\tau}^2$ the pulsar average power

(over the pulsar period) and γ_p^2 and ξ_p^2 two power-independent factors which depend on the shape of the power profile and power spectral density of the pulsar, respectively. The spectral shape factor ξ_p^2 is described by the following equation,

$$\xi_p^2 = \frac{\sigma_{p,\tau}^4}{\sigma_n^4} \left(\frac{1}{N_{sb}} \sum_{i=1}^{N_{sb}} \operatorname{SNR}_i^2 \right)^{-1} \quad (5)$$

with SNR_i the signal to noise ratio in the i -th subband. The temporal shape factor γ_p^2 is given by,

$$\gamma_p^2 = \frac{N_{\tau}^2 \sigma_{p,\tau}^4}{\sum_{i=0}^{N_{\tau}-1} |(d\sigma_p^2/d\tau)(iT_s)|^2} \quad (6)$$

which is sensitive to the derivative of the pulsar power profile $\sigma_p^2(t)$: 'peaky' profiles are good for synchronization.

In the previous derivation, the pulsar frequency relative to the spacecraft frame of reference has been assumed known. A more complete study incorporates the estimation of the pulsar frequency. Nevertheless, under a constant speed model, the CRLB of the pulsar phase coincides with the previously derived equation. Moreover, the pulsar phase estimation errors are statistically independent from the pulsar frequency estimation errors and the phase estimation corresponds to the phase in the middle of the integration period.

2.2 X-ray Pulsar Synchronization

The output of an X-ray detector can be modeled as a Poisson point process, with each event a photon arrival. A *stationary* Poisson point process is characterized by its mean event density (ideally, each event has zero duration) as $\lambda = \lim_{T_o \rightarrow \infty} (n(T_o)/T_o)$ with T_o the observation time and $n(T_o)$ the number of events in T_o . X-ray pulsars are characterized by a *cyclo-stationary* (periodic) probability density of emission. The addition of two independent Poisson point processes (X-Ray pulsar single photon arrival and Galaxy X-Ray Background) has an equivalent arrival density $\lambda(t) = \lambda_p(t) + \lambda_n$, with $\lambda_p(t) \ll \lambda_n$ the periodic pulsar arrival density and λ_n the stationary background contribution. The time resolution of the X-Ray detector is such that if more than one photon arrives within a given time-bin, it is detected as a single photon. For $\lambda(t) \simeq \lambda$ within a given time bin, the probability that k Poisson points occur in an interval of time T_b is provided by the relationship,

$$\operatorname{Pr}(k) = \frac{(\lambda T_b)^k}{k!} \cdot e^{-\lambda T_b}$$

Hence, we have to distinguish two cases: (a) no photon has arrived in a given time-bin of duration T_b ; (b) more than one photon has arrived in the given time-bin. Their respective probabilities are,

$$P_0 = e^{-\lambda T_b}, \quad P_1 = 1 - e^{-\lambda T_b} \quad (7)$$

Detection is modeled by an indicator function I_{n_1, n_2} that asserts whether events are detected in the time bin $t_{n_1, n_2} = n_1 T + n_2 T_b = (n_1 N_b + n_2) T_b$, with N_b the number of bins per period and T_b the detector time resolution. Then, P_0 and P_1 in (7) are expressed from,

$$\begin{aligned} P_0(n_2 T_b) &= \Pr\{I_{n_1, n_2} = 0\} = e^{-\lambda(n_2 T_b) T_b} \\ P_1(n_2 T_b) &= \Pr\{I_{n_1, n_2} = 1\} = 1 - P_0(n_2 T_b) \end{aligned}$$

From the above signal model, the ML criterion consists in maximizing the log-likelihood function $\ln p_1(I|\tau)$ with respect to the pulsar phase τ . Establishing the following probability estimate,

$$\hat{P}_1(n_2) = \frac{1}{N_I} \sum_{n_1=0}^{N_I-1} I_{n_1, n_2} \quad (8)$$

with $\hat{P}_0(n_2) = 1 - \hat{P}_1(n_2)$, the phase estimate is obtained as the maximization of $J(\tau)$,

$$J(\tau) = C_1 + N_I \sum_{n_2=0}^{N_b-1} \hat{P}_1(n_2) P(n_2 T_b - \tau) \quad (9)$$

where $C_1 = -N_I T_b \sum_{n_2=0}^{N_b-1} \lambda(n_2 T_b - \tau)$ is independent of τ provided that the bandwidth of $\lambda_p(t)$ is less than $0.5/T_b$ and, $P(t)$ is defined as

$$P(t) = \ln \left(e^{T_b \lambda(t)} - 1 \right) \quad (10)$$

The CRLB corresponding to this approximate ML formulation results in,

$$\begin{aligned} \text{CRLB}_{\text{low}}^{-1}(\tau) &= \frac{N_I T_b^2}{e^{T_b \lambda_n} - 1} \sum_{n_2=0}^{N_b-1} \lambda_p'^2(n_2 T_b) \\ &\simeq \frac{N_I T_b}{e^{T_b \lambda_n} - 1} \int_0^T \lambda_p'^2(t) dt \leq N_I \lambda_n^{-1} \int_0^T \lambda_p'^2(t) dt \end{aligned} \quad (11)$$

with T_b the duration of the X-ray detector's time bin, N_b the number of bins per pulsar period, N_I the number of integrated pulsar periods and $\lambda_p'(t) = d\lambda_p(t)/dt$.

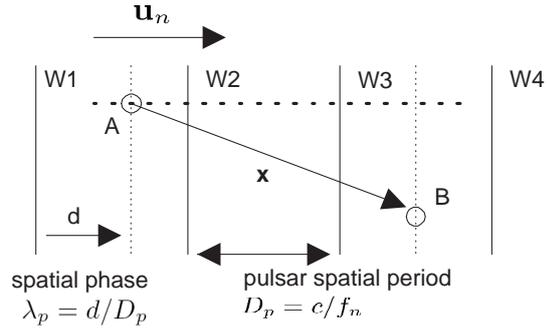


FIGURE 5. Geometrical interpretation of the phase estimate corresponding to time $t = T'_0 + (t - T'_0)$ as a snapshot depicting the pulsar wavefronts (W) as reference. The correction with respect to the reference position (Solar System Barycenter: point A) is represented by \mathbf{x} . The source of ambiguity is observed as an integer number of spatial pulsar periods D_p in the direction \mathbf{u}_n of the pulsar.

2.3 Pulsar Based Navigation

The extreme frequency stability of pulsars allows to predict the phase evolution of the pulsar signal in terms of barycentric time anywhere in the solar system using: (a) a previous estimate of the phase evolution for a reference time T_0 , (b) an estimate of the pulse frequency and (c) its multiple derivatives. The normalized barycentric phase evolution can be expressed as follows,

$$\begin{aligned} \Phi_k^{\text{SSB}}(t_o) &= [\Phi_k^{\text{SSB}}(T_0) + f_k \cdot (t_o - T_0) + \\ &+ \sum_{m=2}^M \frac{f_k^{(m)} \cdot (t_o - T_0)^m}{m!}]_w \end{aligned} \quad (12)$$

with t_o the current unknown time when phase is observed, $\Phi_k^{\text{SSB}}(t_o)$ the phase of the k -th pulsar at t_o , $f_k = 1/T_k$ the known frequency of the k -th pulsar, with T_k its period, $f_k^{(m)}$ its m -th known derivative and $[\cdot]_w$ the phase wrapping operation defined as,

$$0 \leq [\Phi]_w = \Phi + m_\Phi < 1 \quad (13)$$

with m_Φ the integer satisfying the previous condition. For t_o close to the time reference T_0 , we can consider only the first term in the Taylor series (12). Hence, at position \mathbf{x} w.r.t the SSB, phase observations become,

$$\begin{aligned} \hat{\Phi}_n^x(t_o) &= [\Phi_k^{\text{SSB}}(t_o - \tau_k) + w_{\Phi_k}]_w \\ &= [\Phi_k^{\text{SSB}}(T_0) + f_k \cdot (t_o - \tau_k - T_0) + w_{\Phi_k}]_w \end{aligned} \quad (14)$$

with $\widehat{\Phi}_n^{\mathbf{x}}(t_o)$ the phase observation associated with the k -th pulsar at time t_o (the superindex \mathbf{x} is used to denote a spacecraft measurement), τ_k the phase delay w.r.t the SSB and w_{Φ_k} an additive Gaussian noise term associated with the estimation of $\widehat{\Phi}_n^{\mathbf{x}}(t)$. The delay τ_k w.r.t. the SSB depends linearly on the spacecraft position \mathbf{x} as depicted in figure (5): $\tau_k = \frac{\mathbf{u}_k^T \mathbf{x}}{c}$, with \mathbf{u}_k a unitary vector from the SSB to the pulsar (note that \mathbf{u}_k does not depend on \mathbf{x}). Using this definition of τ_k in (15), $\widehat{\Phi}_n^{\mathbf{x}}(t_o) =$

$$\left[\Phi_k^{\text{SSB}}(T_0) + f_k \cdot \left(\Delta t_o - \frac{\mathbf{u}_k^T \mathbf{x}}{c} \right) + w_{\Phi_k} \right]_w$$

with $\Delta t_o = t_o - T_0$. Applying the wrapping operation in (13), $\lambda_k = \widehat{\Phi}_n^{\mathbf{x}}(t) - \Phi_k^{\text{SSB}}(T_0) =$

$$f_k \Delta t_o - \frac{f_k \mathbf{u}_k^T \mathbf{x}}{c} + m_k + w_{\Phi_k} \quad (15)$$

with λ_k the modified phase observation and m_k the unknown number of phase-cycles associated with the k -th pulsar. In vector notation,

$$\boldsymbol{\lambda} = \mathbf{U} \tilde{\mathbf{x}} + \mathbf{m} + \mathbf{w}_\lambda \quad (16)$$

where the following column vectors are defined,

$$\begin{aligned} \boldsymbol{\lambda} &= [\lambda_1; \dots; \lambda_K] \\ \tilde{\mathbf{x}} &= [\mathbf{x}; \Delta t_o] \\ \mathbf{m} &= [m_1; \dots; m_K] \\ \mathbf{w}_\lambda &= [w_{\Phi_1}; \dots; w_{\Phi_K}] \end{aligned}$$

and \mathbf{U} is defined as a $K \times (3 + 1)$ matrix,

$$\mathbf{U} = [\mathbf{U}_0, \mathbf{f}] = \begin{bmatrix} -\frac{f_1}{c} \mathbf{u}_1^T & f_1 \\ \vdots & \vdots \\ -\frac{f_K}{c} \mathbf{u}_K^T & f_K \end{bmatrix} \quad (17)$$

with \mathbf{f} the last column of \mathbf{U} , associated with the time variable. Note in (16) that the observation vector $\boldsymbol{\lambda}$ depends linearly on the unknown parameters $\tilde{\mathbf{x}}$ but is corrupted by an unknown integer vector \mathbf{m} . The determination of \mathbf{m} constitutes the ambiguity problem and it is a detection rather than an estimation problem.

ML Ambiguity Resolution

In the signal model shown in (16), we have two unknowns: the spacecraft's position and time estimate $\tilde{\mathbf{x}}$

and the unknown integer vector \mathbf{m} . The estimation of \mathbf{m} is the most difficult part since its support is the integer numbers. We define $\mathbb{M}[\mathbf{v}, \mathbf{M}] = \mathbf{v}^H \mathbf{M} \mathbf{v}$ for a dummy vector \mathbf{v} , with \mathbf{M} a non-negative definite matrix. Then, ambiguity resolution becomes,

$$(\tilde{\mathbf{x}}, \hat{\mathbf{m}}) = \arg \min_{\tilde{\mathbf{x}}, \mathbf{m} \in \mathbb{Z}^K} [\mathbb{M}(\boldsymbol{\lambda} - \mathbf{m} - \mathbf{U} \tilde{\mathbf{x}}, \mathbf{R}_\lambda^{-1})] \quad (18)$$

where it is assumed that $\mathbf{w}_\lambda \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_\lambda)$ (Gaussian distributed with mean $\mathbf{0}$ and correlation matrix \mathbf{R}_λ). We adopt here the Conditional ML approach (CML) where we calculate an estimate of $\tilde{\mathbf{x}}$ dependent on a dummy \mathbf{m} ,

$$\begin{aligned} \hat{\tilde{\mathbf{x}}}(\mathbf{m}) &= \arg \min_{\tilde{\mathbf{x}}} [\mathbb{M}(\boldsymbol{\lambda} - \mathbf{m} - \mathbf{U} \tilde{\mathbf{x}}, \mathbf{R}_\lambda^{-1})] \\ &= (\mathbf{U}^T \mathbf{R}_\lambda^{-1} \mathbf{U})^{-1} \mathbf{U}^T \mathbf{R}_\lambda^{-1} \cdot (\boldsymbol{\lambda} - \mathbf{m}) \\ &= \mathbf{U}^\# (\boldsymbol{\lambda} - \mathbf{m}) \end{aligned} \quad (19)$$

with $\mathbf{U}^\#$ implicitly defined. This estimate is plugged back into (18) to derive an estimate for \mathbf{m} as,

$$\hat{\mathbf{m}} = \arg \min_{\mathbf{m} \in \mathbb{Z}^K} [\mathbb{M}(\boldsymbol{\lambda} - \mathbf{m} - \mathbf{U} \hat{\tilde{\mathbf{x}}}(\mathbf{m}), \mathbf{R}_\lambda^{-1})]$$

This constitutes an integer weighted least squares problem over an infinite integer lattice with cost function $\psi_1(\mathbf{m}) =$

$$\begin{aligned} &= \mathbb{M}(\boldsymbol{\lambda} - \mathbf{m} - \mathbf{U} \hat{\tilde{\mathbf{x}}}(\mathbf{m}), \mathbf{R}_\lambda^{-1}) \\ &= (\boldsymbol{\lambda} - \mathbf{m})^T \mathbf{R}_\lambda^{-1} (\mathbf{I} - \mathbf{U} \mathbf{U}^\#) (\boldsymbol{\lambda} - \mathbf{m}) \end{aligned} \quad (20)$$

The hereto described technique for ambiguity resolution constitutes a cold-start scheme with no a priori information on the aircraft's position, where $\psi_1(\mathbf{m})$ requires minimization over a wide region of the multi-dimensional integer grid $\mathbf{m} \in \mathbb{Z}^K$. Note though, that as $\mathbf{M} = \mathbf{R}_\lambda^{-1} (\mathbf{I} - \mathbf{U} \mathbf{U}^\#)$ is rank-deficient, there may exist a wide (possibly infinite) range of values for \mathbf{m} almost parallel to the null subspace of \mathbf{M} , which yield a very low (and therefore likely) value of $\psi_1(\mathbf{m})$. Moreover, each of these solutions may correspond to very far locations which are difficult to discriminate (similar values of $\psi_1(\mathbf{m})$) and might easily give rise to an ambiguity resolution error. This problem is addressed next with the incorporation of side information from previous estimations.

Bayesian Ambiguity Resolution

We model the observation vector $\boldsymbol{\lambda}$ (introduced in (16)) and the a priori information about $\tilde{\mathbf{x}}$ as Gaussian ran-

dom variables,

$$\begin{aligned}\lambda &\sim \mathcal{N}(\mathbf{U}\tilde{\mathbf{x}} + \mathbf{m}, \mathbf{R}_\lambda) \\ \tilde{\mathbf{x}} &\sim \mathcal{N}(\tilde{\mathbf{x}}_o, \mathbf{C}_{\tilde{\mathbf{x}}_o})\end{aligned}\quad (21)$$

Now, after some mathematical operations, we can obtain the unconditioned PDF of λ as,

$$\begin{aligned}p(\lambda) &= \int p(\lambda|\tilde{\mathbf{x}})p(\tilde{\mathbf{x}})d\tilde{\mathbf{x}} \\ &= p_0 \cdot e^{-\frac{1}{2}\mathbb{M}(\lambda - \mathbf{m} - \mathbf{U}\tilde{\mathbf{x}}_o, [\mathbf{R}_\lambda + \mathbf{U}\mathbf{C}_{\tilde{\mathbf{x}}_o}\mathbf{U}^T]^{-1})}\end{aligned}\quad (22)$$

The Bayesian estimate of \mathbf{m} becomes $\hat{\mathbf{m}}_B = \arg \max_{\mathbf{m} \in \mathbb{Z}^k} p(\lambda) =$

$$\arg \min_{\mathbf{m} \in \mathbb{Z}^k} \mathbb{M}(\lambda - \mathbf{m} - \mathbf{U}\tilde{\mathbf{x}}_o, [\mathbf{R}_\lambda + \mathbf{U}\mathbf{C}_{\tilde{\mathbf{x}}_o}\mathbf{U}^T]^{-1}) \quad (23)$$

Although this new cost function is different from $\psi_1(\mathbf{m})$, it can be proved that,

$$\hat{\mathbf{m}}_B = \arg \min_{\mathbf{m} \in \mathbb{Z}^k} (\psi_1(\mathbf{m}) + \psi_2(\mathbf{m})) \quad (24)$$

where $\psi_1(\mathbf{m})$ is defined in (20) and the new term $\psi_2(\mathbf{m})$ is defined as $\psi_2(\mathbf{m}) =$

$$= \mathbb{M}\left(\hat{\tilde{\mathbf{x}}}(\mathbf{m}) - \tilde{\mathbf{x}}_o, \left(\mathbf{C}_{\tilde{\mathbf{x}}_o} + (\mathbf{U}^T \mathbf{R}_\lambda^{-1} \mathbf{U})^{-1}\right)^{-1}\right) \quad (25)$$

with $\hat{\tilde{\mathbf{x}}}(\mathbf{m})$ defined in (19). We observe that a priori information has been expressed as the correction term $\psi_2(\mathbf{m})$ to the original cost function, which limits the integer search for \mathbf{m} .

For reasons of space, we have not included the covariance matrix of the position and time errors, which depends on the angular position of the pulsars. Rather we will be using the equivalent error obtained from c (speed of light) times the timing error.

3 Conclusions

As far as signal processing is concerned, the possibility of achieving position accuracies below 10^6 meters has been validated. Nevertheless, the true limitation of pulsar-based navigation has been found to be the required instrumentation. Very few radio pulsars can deliver a sufficiently high SNR with a small antenna at reasonable integration times (few minutes) under the best possible conditions. Results have been obtained for a 10 m^2 antenna, assuming perfect cancellation of radio-frequency noise from the Sun and/or nearby planets,

a constant speed model during integration, perfect antenna pointing (attitude control) and simultaneous pulsar observations. The technological effort applies both to the antenna and the receiver bandwidth. It has been shown that increasing the bandwidth of the receiver leads to better accuracy. A technological limit of 200 MHz at a central frequency of 1 GHz has been considered. Thus, front-end digital signal processing is also performing at high rate, with power consumption a critical factor.

A relaxation of the technological constraints considered in this study (smaller antennas, sequential versus simultaneous observation of different pulsars, smaller receiver bandwidths, provision for implementation loss and safety margins) is possible at the expense of much longer integration times. A reduction in the antenna area by a factor r must be compensated for with an increase in the integration time-bandwidth product of $1/r^2$. An integration time of many hours would be required to extract the faintest but more precise millisecond pulsars from noise. For reduced-size antennas, the constant speed model will fail due to the necessity of estimating higher order derivatives of position. The highest technological impact is the issue between simultaneous or sequential observation of pulsars. The former requires as many antennas as observed pulsars but provides the fastest evolution of positioning accuracy versus time. The latter would further increase latency, the minimum time before a position estimate is produced (probably by a factor equal to the minimum number of required pulsars), and affect the ambiguity resolution algorithm. The smoothness of the spacecraft trajectory is believed to be more critical in the case of sequential observation.

The study of X-ray pulsars predicts longer integration times as arrival rates in the order of only 90 photons per hour have been reported for the usable X-ray pulsars, yielding over one day of integration (taking the ROSAT detector as a baseline). The number of available X-ray pulsars has been found to be much lower than for radio pulsars. Other pulsars with much higher flux densities as the Crab pulsar and improved X-ray detectors would drastically reduce the integration times, but glitch-resilient signal processing should have to be devised and evaluated.

Summary: the authors believe that the complexity involved for autonomous positioning at the spacecraft is rather involved. Although pulsar navigation has been shown to be theoretically possible, issues of complexity and latency in obtaining position estimates cannot be met for *small* spacecraft capable of *fast* autonomous

Pulsar	SNR (dB)	Q (dB)	T_0 (min.)	σ_x (m.)	T_{50} (ms.)
B1937+21	-55.6	8.90	40.13	5860	0.1
B0736-40	-50.2	-10.72	4.02	1699492	29
B1451-68	-50.0	-5.80	2.25	732540	12.5
B0950+08	-48.9	-3.00	1.07	556730	9.5
B0329+54	-45.2	5.32	1.19	77681	6.6

TABLE 1. Minimum Integration times for the best pulsars, using an effective antenna area $A_e = 10\text{m}^2$ and the Gaussian pulse shape model. At least 100 pulsar periods have been considered in all cases. The quality factor Q is a measure of the relative goodness of each pulsar for timing estimation and T_{50} the pulse duration at half power.

positioning: such a navigation system appears more suitable for large spacecraft or for hub stations providing telecommunication and location services to smaller spacecraft. Nevertheless, gains can still be expected from more advanced signal processing, detector technology and combination with other navigation schemes.

4 Vision

Technological requirements for pulsar navigation have been shown to be demanding: radio pulsars require large antennas while X-ray detectors, although much less bulky, have a limited life depending on their gas supply (apart from the faintness of usable X-ray pulsars). Although pulsars are extremely stable, effects such as glitches or pulsar proper motion somehow corrupt the 'neatness' of this timing reference if long-term missions are envisaged. The vision of the authors is more toward a carefully engineered and controlled set of timing references, which, other than using naturally available sources, are tailored to the navigation problem. From our study of radio- and X-ray pulsars, we propose the following approaches:

4.1 Laser-based Navigation

It seems that the bulkiness of the antenna system, due to the extremely weak pulsar signal, would not be required if a network of laser signals were used to transmit carefully designed positioning information over the solar system (a solar-system-scale GPS). New generations of compact and efficient single-photon detectors combined with signal processing could be used on board spacecraft (note that in contrast to X-ray detectors, no gas supply is needed). The positioning algorithms would be more complex as the angular position of navigation

beacons with respect to spacecraft would not be constant over the Solar System (which is the usual case) and it would be definitively more expensive in order to maintain this positioning network. For this approach to be justifiable, a suite of applications where autonomous navigation is necessary would be required. Note that the use of lasers for satellite and interplanetary communications is already under serious consideration.

4.2 Swarm navigation

Distributed signal detection and processing could be carried out by a satellite swarm. Depending on the swarm size, signals picked up by each swarm element could be combined to increase the quality with which pulsar (or maybe artificial beacon) signals are received. In the radio pulsar scenario, the difficulty associated with large antennas might be circumvented as a virtual set of distributed antennas with the same aggregate area could be implemented. In fact, a swarm should be considered a distributed array of detection and computation elements. The problem would be here on how to communicate and perform signal acquisition and navigation calculations over the satellite swarm: the swarm configuration may have to be estimated (we cannot predict the complexity of this operation), as well as a common virtual clock. It would be important to determine what type of information should be exchanged between different members of the swarm to minimize the cost of inter-element communication. Collaborative schemes between elements of the swarm for performing communication (transmission/reception) operations with Earth appears to be also an interesting field, as important as navigation itself.

Acknowledgments

European Space Agency (ESA), project ARIADNA 03/4202 "Feasibility Study of a Spacecraft Navigation System relying on Pulsar Timing Information" April-June 2004.

References

- [1] W. Becker and J. Trümper. The X-ray luminosity of rotation-powered neutron stars. *Astron. Astrophys.*, 326:682–691, 1997.

- [2] W. A. Gardner. *Introduction to Random Processes: With Applications to Signals and Systems*. McGraw-Hill, 1990.
- [3] G. Hobbs, A. Faulkner, I. Stairs, F. Camilo, R. Manchester, A. Lyne, M. Kramer, N. D'Amico, V. Kaspi, A. Possenti, et al. The Parkes multi-beam pulsar survey-IV. Discovery of 180 pulsars and parameters for 281 previously known pulsars. *Monthly Notices of the Royal Astronomical Society*, 352(4):1439–1472, 2004.
- [4] S. M. Kay. *Fundamentals of statistical signal processing: estimation theory*. 1993.
- [5] A. Lyne and F. Graham-Smith. *Pulsar astronomy*. Cambridge Univ. Press, 1998.
- [6] PSRCAT. Atnf pulsar navigation. <http://www.atnf.csiro.au/research/pulsar>, 2004.
- [7] J. Sala, A. Urruela, X. Villares, R. Estalella, and J. M. Paredes. Feasibility study for a spacecraft navigation system relying on pulsar timing information. Technical Report 03-4202, European Space Agency, the Advanced Concepts Team, 2004. Available on line at www.esa.int/act.
- [8] N. E. White, F. Nagase, and A. N. Parmar. X-ray Binaries. In W. G. H. Lewin and J. van Paradijs, editors, *Cambridge University Press*. Cambridge, MA, 1995.



Acta Futura 3 (2009) 125-130

DOI: 10.2420/AF03.2008.102

**Acta
Futura**

Insect navigation and path finding

TOBIAS SEIDL*

Advanced Concepts Team, ESA/ESTEC, Keplerlaan 1, 2201 AN Noordwijk, The Netherlands

Abstract. Insects have a brain weighing about a tenth of a milligram. Nevertheless some insect species exhibit amazing performance in finding their daily paths when looking for food or shelter. Some species have even found their ecological niche in being excellent navigators and hence can survive in extreme habitats. One remarkable and well studied model is the Saharan desert ant *Cataglyphis fortis* that outruns its competitors by performing egocentric navigation. By using skylight cues as a compass and counting steps it is able to find its way without using external visual cues. In comparison, the Australian desert ant *Melophorus bagoti* employs route learning strategies, where it visually learns and in tests recalls every point of their route. In studying the insects' strategies we can learn a great deal on how little information can be used to perform a navigational task. Also the way of how information is processed within such a tiny brain is intriguing. The conclusions drawn from this research are nowadays not only used to understand human behavior but find their way into technical design.

1 Introduction

Evolutionarily challenged organisms optimize among others in terms of energy expenditure (e.g. avoiding detours), information processing (e.g. energy, reliability),

safety (e.g. risk avoidance), failure avoidance (e.g. simplicity, redundancy, failure tolerance). While the qualitative performance is similar - field of neuroethology (behavior associated with sensory information), to outline the lessons learnt on insect navigation so far and finally present some rough ideas for space related technical application.



FIGURE 1. Desert ant *Cataglyphis fortis* initiating a foraging run.

1.1 Insect cognition

Animal orientation can be differentiated into two big fields: While migration is dealing with the relocation of an animal - sometimes over many generations - from

*E-mail address: tseidl@web.de

one place to another, (ii) navigation denotes the systematic return to a previously left point of reference. Both types are found in different taxa, ranging from bacteria to whales. In the following I will focus on the central place navigation techniques as found in insects for a couple of reasons. The insect model is a very variable one and many different "technical solutions" are realized by them. The system architecture of insects is a rather simple one compared to vertebrate models and in consequence easier to understand: First of all the hardware (i.e. brain) is very small and hence simple. Secondly, the behavioral patterns are reduced to the elementary motivations: defense (fight), maintenance (food intake), and reproduction. Since we deal with social insects, where the queen suppresses the workers' reproductive behavior the discrimination of the two remaining motivations is much simpler than when studying a vertebrate model with its huge variety of motivations and behaviors. In many cases, the food items found during a foraging trip cannot be eaten by the insect itself. In consequence the mission goal becomes a rather abstract one, where the agent-ant has to decide if the item found fulfills the profile set by the super-organism. On top of that, it is technically easy to work with insects and so insect models have become very popular among behavioral scientists. Central place foraging on the other hand is the type of behavior that involves highly elaborated signal processing, both in terms of precision and time, and hence appears to be the technically more interesting system. While a migratory animal will interrupt its behavior once a suited living ground is found, a navigating forager is forced to locate the exact point of reference - usually the nest - in order to fulfill its mission.

1.2 Insects qualify in several aspects as models

When reviewing the panel of this workshop you will find two other contributions on insects as model animals for technically oriented studies and application. In a way, this stresses the fact of the amazing success of the insect Bauplan both for the variability of the exoskeleton and the performance neuronal system. While mechanical aspects will be dealt by Stanislav Gorb, Nicolas Franceschini will introduce the insects' visual autopilot. Using both mechanical and neuronal tools, the insects become able to perform high level tasks - also called insect intelligence or insect cognition - such as navigation. In combining the results of these complementary approaches, we will not only learn more about the biological model, but also may be able to transfer this knowl-

edge to technical design.

2 Mastering egocentric navigation: the Saharan ant *Cataglyphis*

One of the best studied models for insect navigation is the Saharan desert ant *Cataglyphis*. Several *Cataglyphis* species have found their ecological niche in areas that seem devoid of life: Desert areas containing almost no vegetation and even less animal life seem not to provide acceptable living conditions. Nevertheless, once spending some time, the observer will discover numerous rather large (approx 1 cm body length) ants with quite a distinct body shape and a rather fascinating behavior: In this featureless habitat, they run at rather high speeds (up to 1 m/s) in a seemingly determined manner toward an invisible goal at the end of their trajectory. Indeed, the ant will eventually disappear in a small inconspicuous hole in the ground - the entrance to the colony's nest, hosting up to a few thousand individuals. A deeper study will reveal, that emerging from this nest, a cohort of ants will go foraging for food throughout the day, performing one of the most fascinating behaviors in animal kingdom. Every little ant will leave the nest for a distance of up to twenty thousand times their own body length, and search for a single food item - a dead insect for example that succumbed to the atrocious climatic conditions. Once successful, the ant will turn around and in a determined manner steer back home to the nest in a direct line without getting lost. Even in the absence of any visual cue, i.e. landmarks, the ants will successfully locate their nest and immediately initiate another foraging run. This behavior is somewhat different to the commonly known routes that can be observed in the majority on the European ant-species. Wood ants for example establish a route to a feeding site, i.e. a dead mouse, by placing olfactory marks along the way. Other ants will follow this route and easily find the food source. Each of the ants will add its olfactory share to the route and in consequence, the route will be reinforced and even attract more ants, until the source is getting exploited. Now this approach does not work with desert ants for several reasons. Firstly, the light sand in the desert has little cohesion and hence the route marked onto it will be carried away easily by wind. If not the wind, the dry and hot climate will - secondly - evaporate the pheromones too fast in order to allow for a stable marking. And - thirdly - there are almost not abundant food sources available that would require

or justify establishing a route. Food items are sparsely distributed and hence can usually be carried away by a single ant. Passing on the knowledge of a feeding site to a nest mate would not contribute to the colony's success. On the other hand, each ant that can be found outside the nest follows the very same programme of searching for food and then safely and quickly navigating home. Hence desert ants are an ideal study model for (biological) autonomous agents.

2.1 The compass

Cataglyphis ants navigate egocentrically, i.e. without using external visual cues. This implies that they constantly monitor their own movements and update the knowledge of their position in respect to the nest, the home vector. Path integration - as the basis of vector navigation - requires knowledge of the direction, distance and inclination of the current segment of the path in order to calculate the home-vector. The by far best studied element of desert ant navigation is the compass mechanism with which the ants determine rotation around the vertical axis. The work on the skylight compass of insects started with von Frisch's [3] Nobel-awarded discovery that bees use polarized light to determine flight directions. In short, many insects, including *Cataglyphis* have a part of their compound eye specifically adapted for compass tasks. The dorsal rim region, i.e. a small part on the very top of the head, has eyes that are sensitive to polarized UV-light of a preferred angle arranged in a fan shape configuration [12, 2, 10]. Theoretical considerations propose the existence of three integrator neurons that are tuned to different directions and subsequently pass on the information of direction to so called (and proposed) compass neurons. However, there are some apparent constraints when exploiting the polarization of the sky for directional purposes. (i) Light, has to be present. This means only during day and nights with full moon, the quality of the signal is good enough for the ants. (ii) Visibility to the sky has to be granted to a degree of about 20% that is free of clouds. (iii) During the day the sun turns at changing speed. The compensation of the rotation of the sky-pattern has to be performed accordingly. (iv) During the year, the so called ephemeris function changes its position, and hence, the pre-programmed compass has to be adaptive to that as well. (v) Finally, the ephemeris on the south hemisphere is flipped. Even more so if the colony is located between the two tropics and hence the December worker has to cope with a flipped version of

the curve, the July worker is confronted with - both potentially being clones. The relationship between genetically inherited and deductively learnt information on the current shape of the ever changing ephemeris function is addressed in currently ongoing studies.

2.2 The odometer

In contrary to the compass mechanism, it has long been an open question on how desert ants measure distances run. Several hypotheses had been posted, i.e. monitoring the energy spent, or the time spent to cover a certain distance. Both were discarded for various reasons and hence only exploiting visual cues and proprioceptive ones (i.e. own movement) remained. While flying insects such as bees use optic flow for optometry, this factor plays only a minor role in running animals (16%, [8, 7]). A series of studies finally demonstrated, that distance estimation is performed by a pedometer, i.e. a step integrator. In order to proof this hypothesis, [13] trained ants to forage to a feeder. Once arrived at the feeder, the ants' leg lengths were manipulated by microsurgery operations, leading to longer (stilts) and shorter leg lengths (stumps). Once released in a test array, the ants on stilts overestimated the way home and performed their search pattern behind the position of the virtual nest, while the stumps-ants under estimated the distance and looked for the nest entrance at an earlier point of their search. However, as step length changes with speed of locomotion, the ants do not only sum up the number of steps, but also incorporate the step length. Studies on the kinematics of running ants have not revealed the sensory mechanism responsible, but point towards force sensors in the animals' leg, i.e. muscular tension receptors or strain sensing campaniform sensilla in the exoskeleton.

2.3 Integrating inclined paths

The initiator for Wittlinger's studies was an experiment performed some years earlier and that revealed an interesting feat: When outbound ants, i.e. on the way from the nest to the feeder, were running on a corrugated path, i.e. crossing a series of several artificial hills, they performed much longer path than the actual bee-line between nest and feeder measures. Once equipped with a food item and placed into a flat channel, the ants performed the typical nest search after they had covered the shorter of the two distances [14, 15]. Obviously, they integrated the inclination of the paths run

and calculated the ground projection, which was the basis for the home vector. This performance can be explained only by means of proprioceptive monitoring of the inclined surface during walking. All other hypotheses such as visual integration of inclination, or energy and time measures fall to fundamental flaws. However, the exact mechanisms of monitoring inclination for means of path integration has not been found so far. Again, force feedback during walking currently seems the most favored explanation.

2.4 Coping with errors

A typical foraging run of a desert ant can involve about 20 thousand single steps which are integrated to continuously update the information on the shortest way back to the nest - home vector. It is evident, that even slightest errors may severely deviate the animal from its track and induce fatal consequences. In the following I want to outline some of the techniques used by the ants' navigational toolkit in order to maintain a satisfactory success rate.

Systematic search

Once the ant has performed its run back home and the home vector reaches a state close to zero, it will switch from the linear path to a series of loops with increasing diameter with the cross point centered on the position of the expected nest entrance. With this systematic loops, the ant will search for the nest entrance in the vicinity with an intensity following a statistic distribution.

Linear and angular undershooting

In linear test set-up, the estimated nest entrance moves toward the feeder with increasing homing distance. In angular outbound paths, the return angle is systematically pointing not toward the virtual nest but toward an estimated nest entrance moved along the outbound path. An ant that has performed several foraging runs, will know the visual features of along the usually taken path better than those "behind" the nest (ants show high sector fidelity). Once a known structure is identified, the previously lost nest entrance is easier to be located.

Resetting the vector

During each foraging run, the ant will incrementally accumulate errors on the position of the nest. The reference point with the highest significance to the animal

is the nest. Hence, only in the nest eventual (or induced) errors on the vector will be deleted. Every ant leaving the nest for a foraging run will be in an initial zero-vector state.

Exploiting external cues

Even in the most homogeneous environment once in a while, unique structures appear and hence can be exploited by the ant for navigational purposes. May this be visual landmarks such as tussocks or stones, tactile landmarks [9] such as the roughness of the ground or even olfactory landmarks not connected with food or nest scent. These landmarks may temporarily override the path integrator, but not reset it.

3 Coping with the labyrinth: the *Melophorus*-approach

Cataglyphis is not the only genus of ants that occupy the niche of hot and dry habitats. In the south of Africa we find *Ocymyrmex velox* and in the Australian outback *Melophorus bagoti*. The later one exhibits a quite fascinating approach to cope with the peculiarities of its habitat: Numerous and densely placed tussocks form rather a huge labyrinth than a free and uniform are such as we can find it in the Sahara [6]. In consequence, finding its way home in a time efficient manner is not only a matter of getting direction and distance right, but also avoiding dead ends. In addition to that, potential predators can hide much easier and hence a safe route is worthwhile to have. In consequence, *Melophorus* ants establish routes to an installed feeder and maintain these routes rather conservatively. On top of that, the outbound route and the inbound route do not have to be identical. On the other side a displaced ant will initiate search loops and once it found its previously learnt route immediately recognize it - if it is the right one - and continue to follow the route without hesitation. The exception to be made is, that an ant on outbound motivation will not recognize its inbound route and vice versa. Now the most intriguing question is, on how an ant with a brain of approximately 0.1 mg will be able to remember every section of a sequence of several thousands of "snapshots" and quickly retrieve the right match in the instant of being confronted with it. Obviously, it cannot be a simple hard-drive with a series of bitmaps that are constantly screened. Even more so in the sight of recent results that showed that ants can learn and retrieve up to three separate routes. Theoretical studies on

feature extraction and localization try to shed light onto this so far not understood mechanism.

4 Insect intelligence in the space context

The the selection pressure acting upon the foraging abilities of the species introduced in the present account has led to an optimisation of the ants' major survival feature: The navigational toolkit. Several different systems act in parallel in order to allow for the ant to perform successful navigation. Although it is impossible to determine what "Evolution had in mind" we can identify a number of basic demands that appear to be set during the evolutionary development. An ant brain is rather small and computing power as well as data storage is highly restricted for several reasons. Both energy constraints [5, 1] as well the general Bauplan do not allow for indefinite expansion of brain tissue. Hence incoming data has to be highly reduced to the relevant and only the relevant bits. Extracting behaviorally interesting features from the incoming data has to be performed in a real time manner since predators are not known to wait. Additionally, the calculations performed have to be done so in a reliable way, since errors cannot be monitored or corrected during a foraging run. Also recalibration is only known for subsystems such as the compass. The vector is only reset in the colony. On top of all that the ant has to deal with insufficient sensory data. Its knowledge of the world is highly limited and it can be assumed also noisy. The study of these highly evolved and carefully tuned system allows us to draw some conclusions that can not only entertain biologists, but also lead to a systematic application in technological development.

- Each *Cataglyphis*-ant is able to perform foraging runs with 20.000 translational steps including rotations and inclines, integrating iteratively every single vector at a precision sufficient enough to relocate the nest while running on a sometimes slippery ground in a rather shakey manner. Neuronal mechanisms of data fusion, error elimination, and the lot may provide new tools for various applications.
- The architecture of the e-vector analyzers including the subsequent hard-wired data processing leads to a fully functional skylight compass. It is robust enough to work under suboptimal conditions such as almost obscured sky or even during full

moon nights. Still, it is a simple electronic system that has already found its way into a technological demonstrator [4, 11]

- Only little exposition to the sunlight is necessary to determine the current shape of the ephemeris function. Understanding the balance between pre-programmed and acquired knowledge including the minimum input necessary to perform this task may lead to a deeper understanding for adaptive mechanisms in autonomous systems.
- Not all available data is used by the sensory system. Which features of a visual scenery is finally memorized (snapshot) and subsequently compared to the actual scenery? The algorithms that are used by the technological competitor to locate the nest with visual cues could be useful for ones own projects.
- On top of that, the Australian ant, *Melophorus bagoti*: How does such a small and rather unintelligent being memorize and retrieve every single point of its route(s)? Again feature extraction, data compression and comparison algorithms could be an interesting starting point in machine vision projects.
- One of the big questions in animal navigation in general is the dispute between map-like representation or procedural knowledge of the habitat. The current state indicates that the use of maps seems to be rather improbable since simpler models allow to explain all current data. Might autonomous exploration vehicle be better off to use this form of "mapping" their environment?
- One of the most amazing observations is that foraging ants usually fall to predators but only rarely get lost or miscalculate their energy budget. The ant-controlled part of a foraging excursion appears to be rather robust and "mission proof". Examples are the self elimination of inevitable errors or the temporary dominance of the visual navigation system over the egocentric one during nest search. The basis on which information is evaluated and trusted can inspire the development of autonomous, redundant systems in space applications

The lessons that can be learnt from systematic analysis of biological systems is not only a time worthy activity but may also lead to new insights for technical ap-

plications. Biological systems are suboptimal in the respect that they have to follow the genetic history of the Bauplan and cannot design freely in space and material. However, the solutions achieved are sometimes unique and justify a deeper look. On the other side, technological research has the big advantage that the competitor may be analyzed at any time and subsequently imitated if the solution appears to be better. The competitors emanating from nature should not be underestimated.

References

- [1] J. Allman. *Evolving Brains*. WH Freeman & Co, 1999.
- [2] K. Fent. Himmelsorientierung bei der Wüstenameise *Cataglyphis bicolor*: Bedeutung von Komplexaugen und Ocellen. *Zürich, Universität Zürich*, 157, 1985.
- [3] K. v. Frisch. Die Polarisation des Himmelslichts als orientierender Faktor bei den Taenzen der Bienen. *Experientia*, 5:142–148, 1949.
- [4] D. Lambrinos, R. Moller, T. Labhart, R. Pfeifer, and R. Wehner. A mobile robot employing insect strategies for navigation. *Robotics and Autonomous Systems*, 30(1-2):39–64, 2000.
- [5] R. Martin. Relative brain size and basal metabolic rate in terrestrial vertebrates. *Nature*, 293(5827):57–60, 1981.
- [6] B. Muser, S. Sommer, H. Wolf, and R. Wehner. Foraging ecology of the thermophilic Australian desert ant, *Melophorus bagoti*. *Australian Journal of Zoology*, 53(5):301–311, 2007.
- [7] B. Ronacher, K. Gallizzi, S. Wohlgemuth, and R. Wehner. Lateral optic flow does not influence distance estimation in the desert ant *Cataglyphis fortis*, 2000.
- [8] B. Ronacher and R. Wehner. Desert ants *Cataglyphis fortis* use self-induced optic flow to measure distances travelled. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 177(1):21–27, 1995.
- [9] T. Seidl and R. Wehner. Visual and tactile learning of ground structures in desert ants. *Journal of Experimental Biology*, 209(17):3336, 2006.
- [10] R. Wehner. Himmelsnavigation bei Insekten. *Neurophysiologie und Verhalten. Neujahrsbl. Naturforsch. Ges. Zürich*, 184:1–132, 1982.
- [11] R. Wehner. Desert ant navigation: how miniature brains solve complex tasks. *Journal of Comparative Physiology A: Sensory, Neural, and Behavioral Physiology*, 189(8):579–588, 2003.
- [12] R. Wehner and S. Strasser. The POL area of the honey bee's eye: behavioural evidence. *Physiological Entomology*, 10(3):337–349, 1985.
- [13] M. Wittlinger, R. Wehner, and H. Wolf. The Ant Odometer: Stepping on Stilts and Stumps, 2006.
- [14] S. Wohlgemuth, B. Ronacher, and R. Wehner. Ant odometry in the third dimension. *Nature*, 411(6839):795–8, 2001.
- [15] S. Wohlgemuth, B. Ronacher, and R. Wehner. Distance estimation in the third dimension in desert ants. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 188(4):273–281, 2002.

Workshop Conclusions

Speech by Géraldine Naja-Corbin, closing the two day workshop

THE first remark is on the time scale we have been envisaging with this workshop: We are talking of a time span beyond 2030, thus outside the borders of the current programmes and plans which reach approximately until 2020. Choosing this time scale allows going beyond simple extrapolation of current situations and trends in order to envisage real conceptual and technical breakthroughs.

The four panels have allowed us to address the general scientific and technological context, which will influence the future of space in Europe, to envisage expected breakthroughs and their consequences on space systems.

We have learnt with wonder how much nature could inspire us for future technologies and systems and how the world of the living and the world of engineering were now intertwined. On energy, we have seen that gravitational and kinetic energy could be extracted out of orbiting natural satellites. Extracting energy out of the solar wind may be an interesting thing to investigate. Also biomass fuel cells could be an attractive option for the space sector, e.g. ISS or manned exploration missions, since they could solve the issue of waste management and energy re-utilisation. Generally we have received confirmation that the breakthroughs of tomorrow will probably be at the boundary of scientific disciplines and that we must promote multi-disciplinarity and interdisciplinary dialogue.

Now the question is: what are the consequences for us, more modestly in 2008 in ESA?

Alain Dupas presented a global forecast for the trends that will influence our societies' evolution in the long-term future: general technological and societal evolution, international context, and the needs for space (services/infrastructure) arising from public policies and commercial endeavours.

Now I would like us to reflect on the consequences of these global trends on space activities in Europe, and also on possible scenarios for the future of space activities. Reflecting back on the options considered, I see

three scenarios for the use of space in the long-term – by the way, these three scenarios are not mutually exclusive and the final scenario may well be a pondered mix of the three:

- **Scenario 1:** Protection and Security
- **Scenario 2:** Outreach and Expansion
- **Scenario 3:** Rational Use

Scenario 1 is the scenario in which Earth sees itself primarily as exposed to a wide range of threats and dangers. Some of these threats are Earth bound – environmental/climate change, pollution, nuclear proliferation... some are linked to and even coming from space: NEO's/asteroids, debris, solar eruptions, militarisation of space... In this scenario space means are used primarily as a means to protect Earth from dangers (some of which come from space), and Earth is "self-centered"; the space programmes which are promoted are:

- monitoring systems for climate and environment
- early warning systems both for military purposes and for natural disasters purposes
- combination of space systems for the prevention of mitigation of disasters
- space weather systems
- in the longer term, climate control systems?

Scenario 2 is the scenario in which Earth is in an expansion dynamic move and sees the expansion towards outer space as a way to solve its Earthly problems. Space is seen as an end in itself and as a place to conquer, colonise, explore... so here we are looking at the following programmes:

- large scale exploration of the Solar System, both automatic and manned
- installation of manned outposts
- use of extraterrestrial resources
- solar energy from space, solar power systems, either in orbit or on another bodies
- in the longer term, terraforming to create an alternative to Earth?

Scenario 3 is the scenario in which space has no specific part, neither representing a potential danger as in scenario 1 nor an opportunity as in scenario 2. It is just a means and a tool for policies, and to create services and businesses. In this scenario the typical space programmes are:

- integrated services using different space and non-space systems
- a special application could be space tourism.

Two remarks need to be made in this context:

- science (space sciences, Earth sciences and life & physical sciences) is not mentioned specifically in any of the 3 scenarios because it is assumed that science (as is currently performed in ESA's scientific programmes) will continue to be the backbone of European space activities; thus science is actually included in all three scenarios;
- a breakthrough in launchers and particularly in chemical propulsion (even if we have been told that this was unlikely in the medium term future) would probably have an impact on the type of scenario and on the extent of use of space in daily life.

Based on these and considering that in any scenario, our societies will depend on space, the key question for us in ESA is:

What can we do now to prepare for the future? What decisions must be taken now and in the short term future – 5 to 10 years – to ensure that space is used optimally for the future of Europe and of humankind more generally?

We have got some very interesting inputs to these questions, thanks to all participants who have accepted to reply this morning, when we circulated two forms asking which was the most important decision ESA should take, and which was the most important technology ESA should invest in.

In summary, the inputs have been the following:

- **on the most important decision ESA should take,**
 - creating stronger links with the EU and other Agencies was seen as an important item, as well as developing new partnerships;

- an emphasis was placed on innovation and technology support;
- promoting spin-ins and multidisciplinarity;
- increasing visibility and awareness of space;
- particular importance of **risk**: we should go for high-risk approaches, and even dedicate a fixed share of our resources to high-risk technology demonstration projects;
- and then, some original suggestions: create a European Space University or allow for start-ups in space tourism.

• **on the most important technology ESA should invest in:**

- the most important technology area is considered the ones related to launchers and propulsion in general, as this is clearly a prerequisite for space development;
- biotechnologies, in particular bioinformatics, biosensors;
- technologies related to energy: wireless transmission, NPS, solar cells...;
- technologies related to Space Situational Awareness and Near Earth Objects (either to be protected from them, or to mine them!);
- nanotechnologies;
- and finally artificial intelligence.

This could in fact be the closing statement of this panel: we must invest in intelligence, be it natural or artificial, individual or collective. We heard today that systems will become more and more intelligent, in their design and operations. In turn this is what the ACT is about and this is probably the smartest investment ESA could do in the future and for its future: Invest in Intelligence!

Géraldine Naja-Corbin
Head of Institutional Affairs and Strategic Studies
European Space Agency (ESA)