Table of Contents

NanoSatellite Engineering Professional Certification KSF Space Foundation

Nanosatellite Engineering Professional Certificate: NEP is the world's first and only program of curriculum developed by engineering professionals from the KSF Space Foundation that is targeted for universities, startup companies and all of the space industry worldwide to enlighten engineers, teachers and students on construction and development of your very own Nanosatellite. The program will provide essential information on the process from mission development and design all the way to launch and policy regulations.

With three major components that are:

-How to employ nanosats/ (Building a mission)

-How to develop a nanosat/ (Design and build the system)

-How to operate a nanosat/ (Operation, end of life, standards)

There is a list of chapters' users will study and once completed will be able to start their own mission and construction of a NanoSatellite from start to finish with the knowledge acquired from this program.

There are ten chapters that cover each important aspect in designing your nanosatellite that experts have composed for easier learning and applicable knowledge.

Nano Satellite Technologies and Infrastructure

1) Mission development and System Setup
2) Nano-Micro Satellite architectures
3) Component technologies
4) Software development
5) Ground station systems
6) Computing & Attitude control
7) Propulsion
8) Power, Telecommunications & Thermal management
9) Launch opportunities & deployment
10) Regulatory & policy issues

Authors, Reviewers and Contributors

Dr. Kayyali Mohamed, Chairman KSF Space, UK

KSF Space Foundation- UK

Randa Milliron- CEO/Co-Founder- Interoborital Systems-US

Jacob R. Wade- Space Systems Engineer- KSF Space Member- NEP Project Manager- US

Javier Quistiano- Aeronautical & Space Systems Engineer- Mexico

Luis Gerardo Zárate Vilet- Space Systems Engineer- Mexico

Rushi Ghadawala- Space Scientist/ Engineer& Researcher- Canada

Ricardo Colpari- Satellie Technologies- Bolivia

Sundaravathanan Chelleppan- Propulsion Engineer- Malaysia

Elyka Abello- Spacecraft Power Supply Design Engineer - Venezuela

Ravi Kumar- Lead Engineer- Thermal and Fluid Systems- India

Ahmed Abdi- KSF Space Member- Astropreneur- UK

Robert Bernard- KSF Space Member- Aerospace Engineer- Canada

Alok Shrestha- Senior Research Scientist- SSAI- Science Systems and Applications Inc.- US

Colonel Abdel Ghani Toumi- KSF Space Member-

Chapter 1: Nanosatellite Architecture

Nanosatellites are the new and improved method of getting into space. Instead of spending high amounts of money, patrons and small companies now can design and launch multiple Nanosats into orbit for the same price or less than a typical commercial satellite. All satellite missions begin and end in the same place with the mission development being the starting point and orbiting in space as the final point. It is what goes on between those steps that is variant and makes the mission complex. The mission development and NanoSatellite Architecture will be elaborated on, aiding you to start your own specific mission to get into space.

Fundamental Mission Design:

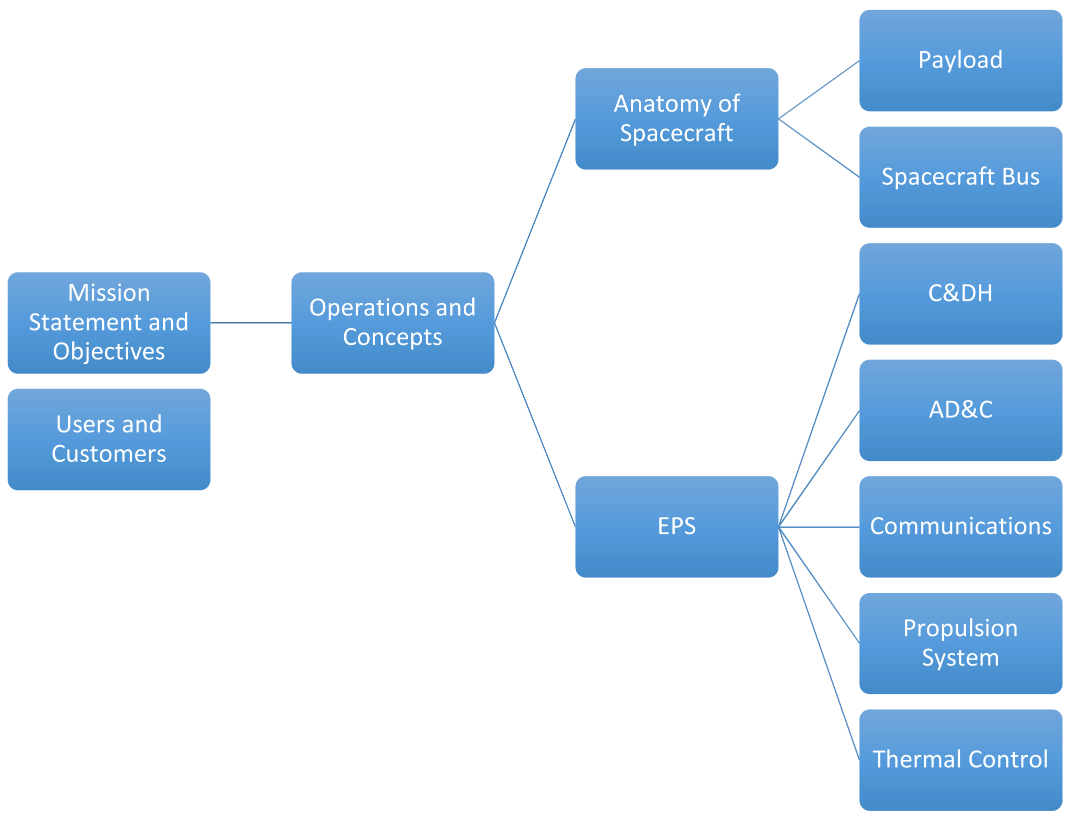


Figure 1.1 (Above) Block Diagram of Fundamental Mission Design Process (KSF SPACE)

1.1 Mission Statement and Objective

The mission statement and objectives are the most crucial and key parts to identify before anything else happens. This process gets the mission in action by identifying the most important aspects and describes what the mission of the NanoSatellite will complete when launched into space.

The mission statement should establish key elements that will be completed, for example: scientific mission, communication link, military use, remote sensing and imagery etc. The mission statement is very important because it will be referred to after each process is finished to ensure the mission will be completed and is still on track. There should be more than one mission statement as well, first is the primary mission statement that will be the most important task at hand.

There also needs to be secondary mission statements that coincide with the primary mission, to have a more well-rounded approach and mission redefined for the customers and users your NanoSat was designed. While developing the mission concept always keep in mind the cost for each subsystem and components, to keep the price in the budget.

Example 1.1

Primary Mission:

NanoSat 135 will analyze the Earth’s atmosphere and collect data for plasma particle concentration and also determine atmospheric measurements in different zones.

Secondary Mission:

Relay plasma particle concentration to research centers.

Measure atmospheric barometric levels at different levels of atmosphere.

1.2 Users and Customers

The customers of your satellite mission are the people and organizations that will benefit and analyze the data and research collected by the satellite mission. It could be a university, commercial industry, military, or personal use for a start-up or organization. Users are the team that will operate the satellite from the ground or in space on the International Space Station. They will be the ones collecting the telemetry communicated to the ground station and interpreted and sent to the customers, ensuring mission success and their data is collected correctly. This is essentially the “who” factor in mission design.

1.3 Space Environment and Risk

The space environment is the harshest most relentless place in the universe known to mankind. When designing a satellite, you have to consider the severe levels of radiation being exposed to the spacecraft and also keep in mind that there are countless of micro meteorites that are constantly bombarding the satellite on orbit. When designing components and sub systems there needs to be rad hardened electronics and materials that are integrated onto the spacecraft. “For small satellites, e.g. cansats and CubeSats, part of the approach is to use COTS components to reduce system whole of life costs. Current electronics from top providers uses these kind of components, even though it is desirable to use rad hard components the price of these devices is considerably higher and most of the times a COTS component are used instead. An example of a rad tolerant processor designed for space is the ATmegaS 128 which has an estimated price of around $670 USD. GomSpace, one of the top providers or CubeSats components uses non-rad hard, nor rad-tolerant devices in its on-board computers such as the Xilinx Zynq 7030 with an estimated price of around $125 USD”

There also needs to be in depth testing to determine the temperature fluctuation in orbit around the Earth from sunlight to shade. Since space is a vacuum you also need to make sure the components and subsystems do not outgas and cause harm to other subsystems onboard the spacecraft. There is also a risk of colliding into space debris and other orbiting satellites that are de-orbiting or not operating anymore. Atomic oxygen also known as rusting plays a role in spacecraft depletion along with the intense levels of high energy photons and electrons from gamma rays, x-rays and ultraviolet rays emitted in the electromagnetic spectrum. These high-energy interactions with your spacecraft will significantly decrease the lifetime of your satellite in orbit, so it is important when determining the mission lifetime and success rate. Space also has electromagnetic effects on electrical components especially in low earth orbit interfering with the Van Allen belts and Earth’s magnetic field.

1.4 Operations and Concepts

Operations and concepts are the key elements of mission success because this determines the works behind the spacecraft and all the long hours out in designing, fabricating and manufacturing, testing and integration, launch and spacecraft operations once in space. This is a vital role in the mission success because you have to define each building block and subsystem in order to achieve your goal. You need to assign engineers, scientists, programmers, machinist, testing equipment and facilities that will be used. Once you assign all the work and who will be completing these tasks there has to be a cost factor involved that stays within the means of your targeted budget established in the mission design. Mission drivers are also an important factor to keep in order like components, the number of satellites, power budget, altitude and orbital parameters, instrument size and also spacecraft size and weight, etc. It is an industry standard to have three separate models of your spacecraft the first being a "3-D model" to model the initial spacecraft concept to fundraisers and supporters of mission. The second is an "engineering model" that is for actual testing and integration of sub systems to test functionality and make sure all systems are online and communicating and have no faults along with environmental testing (thermal vacuum baking and temperature cycling simulating orbit and Residual Gas Analysis testing) along with spacecraft vibration analysis to simulate launch profile on the rocket that will take spacecraft into its final destination in space. The final model is the "flight model" that is the actual satellite that will be launched into space. Operations also includes the ground system set up and link budgets for communications, who will be your launch provider, where and when will the launch be, and final delivery date of your spacecraft. Once these steps have been analyzed and set into place the mission is a go and you can begin to fully design and manufacture your nanosatellite now that the mission is defined and setup for success and remember to keep in mind the mission objectives when diving head first into production.

1.5 Anatomy of Spacecraft

Each section below is a basic overview of each system and will be elaborated in detail in the following chapters.

Payload

The payload is the primary instrument that will perform the science in space or primary mission objective. Deciding and designing the payload is the most important component of the spacecraft because it will be the most crucial aspect in mission success, due to the fact it will be collecting all the scientific data. Every other subsystem will be designed to keep the payload operational and safe on board. The bus will be the spacecraft carrier to space and house all the other subsystems on board. The spacecraft bus must be designed around the payload in order to protect and remain in the form factor restraints of the satellite.

Bus (Satellite Structure)

The spacecraft bus is also important but not as important as the payload, that's why the payload is first and the bus is second in order of design. The bus must meet specific qualifications, size and mass constraints in order to be launched into space because of the limited amount of space in the launcher inside the rocket or launch vehicle, especially with nanosatellites due to the number of spacecraft being taken into space at one time. You wouldn't want to have your satellite get all the way until pre-launch checks and fail the size and mass constraints and not be launched in the rocket.

EPS

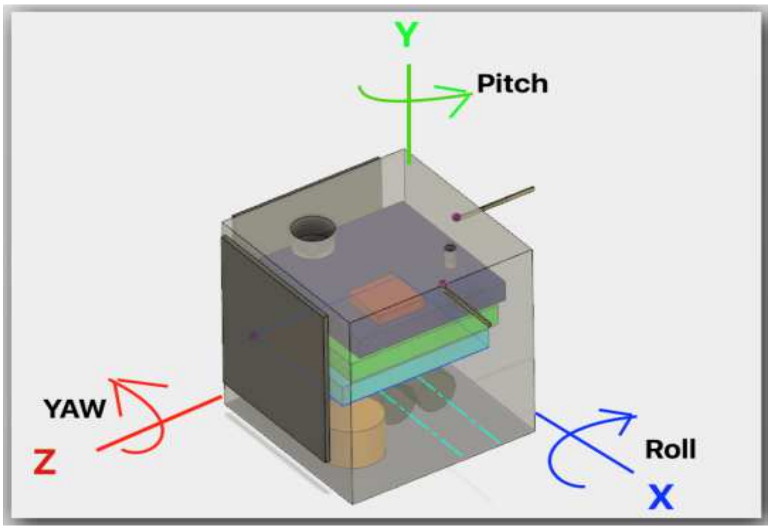
EPS is an acronym for Electrical Power System. This is the subsystem that controls all the current and voltage and power regulations to all systems aboard. There are voltage regulators, op amps, capacitors etc., all integrated into an integrated circuit (IC). The power systems keep the electrical power system and batteries charged but are controlled by the EPS. The payload usually draws most of the current the EPS so it is mainly designed for the payload and all other systems will be designed around what is left after the power budget is constructed for the payload.

C&DH

Control and data handling is the brains and computer on board the spacecraft and does all the computing and data storing and sorting and relays to the earth station at the data packet size and rate determined in C&DH design. This is controlled by a microcontroller and where all the software development will be implemented into your system. The C&DH controls all systems with commands and signals created into the embedded software by software engineers. This is also the method of updating the satellite from earth by re programming the software when systems malfunction or need to be tested for performance and controls the power modes and resets in orbit. So, without the brains there will be not be an operational satellite because this controls all functions in orbit.

AD&C

Attitude Determination and Control, is the subsystem onboard that determines the direction, location, altitude, nadir, zenith and azimuth of the spacecraft. This controls the roll, pitch and yaw, velocity and acceleration etc. of the satellite when orbiting the earth to keep spacecraft from tumbling out of control causing mission to be unsuccessful and data to be useless. The AD&C also controls the velocity and vectors of your satellite so when you need to face the sun for solar panel charging there is a sun pointing vector that will orient the spacecraft to charge. Likewise, when needing to study the earth or deep space or a point source in the galaxy it all depends on the mission needed to be completed. This system will also have control over some other subsystems like propulsion in order to move the spacecraft into a specific location in orbit or to move away from debris etc.



Communications

The communication system is critical to mission success because without being able to communicate to your satellite the mission is a failure. This is also very important in relaying data to Earth stations and other satellites in a constellation surrounding the earth to relay messages to other spacecraft in you mission. There is an onboard radio that transmits and receives data to your ground station and mission operation center to update spacecraft health and system information along with scientific data to pass onto customers that are in need of your satellites mission. There needs to be a link budget determined that defines all the communication details before launch, like the size of your bandwidth, location and size of antenna on earth and on the satellite, along with data rates and data packets that will be pinged to the ground station when fly by in orbit. It is best to communicate with the satellite when it is the highest points in the horizon on Earth due to the noise and background interference from cell towers and day to day life on earth. Radio frequencies are the industry standard for communication due to the wavelength and reliability of data transfer. There is a range of frequencies from 3kHZ to 300GHZ that is primarily used for wireless communicating along with several different bandwidths like L band, S band, and KU band to name a few.

Note: These elements below aren't always integrated in the design of CubeSats. Thermal control requires elements like heat exchangers that requires lots of space leaving no room for payload in small CubeSats e.g. 1, 2, 3 units. Propulsion systems have the same problem than thermal control systems. However due to the importance of propulsion in a satellite, there are research works aiming to reduce the space required to embed small propulsions systems in CubeSats.

Thermal Control

Without thermal control your satellite would either over heat at a rapid pace or freeze in the harsh environment of space temperature modulations. There are many precautions and processes to follow in order to achieve optimal protection of your spacecraft's components and frame. Most subsystems that are subject to overheat will be protected by a heat sync integrated onboard to draw excess heat away from important components. There are also heat resistive materials that are integrated onto the spacecraft that will help protect from overheating and freezing in the extreme thermal cycling while orbiting earth, because the intense fluctuation of temperatures wears on materials and components overtime. Testing each subsystem and the spacecraft as a whole in thermal vacuum chambers is essential to analyze how your system will perform in space before launch.

Propulsion

Last but not least is the propulsion system that is in control of majority of the spacecraft's movements. It is mainly controlled by the AD&C to ensure the spacecraft is on orbit path and in the right orientation if the spacecraft moves out of position. There are many different types of propulsion systems like: cold gas thrusters, electronic ion thrusters, chemical propulsion, and solar sails to name a few. This area has been under increasing development in the past couple pf years especially with spacecraft becoming smaller like nanosatellites. There cannot be a big propulsion tank on board because of the size and mass constraints so there has been development in shrinking propulsion systems as well as advancements in propulsion methods. Propulsion is very important after satellites leave launchers and are in space to ensure the spacecraft moves into the right orbital parameters and altitude and remains there until commanded otherwise.

Works Cited

Pratt, Timothy, Charles W. Bostian, and Jeremy E. Allnutt. Satellite Communications. New York: Wiley, 2003. Print.

Sellers, Jerry Jon., Wiley J. Larson, William J. Astore, Anita Shute, and Dale Gay. Understanding Space: An Introduction to Astronautics. New York: McGraw-Hill, 1994. Print.

Wertz, James Richard. Space Mission Analysis and Design. El Segundo: Microcosm, 2003. Print.

Chapter 2: NanoSatellite/ CubeSat Structure (Spacecraft Bus)

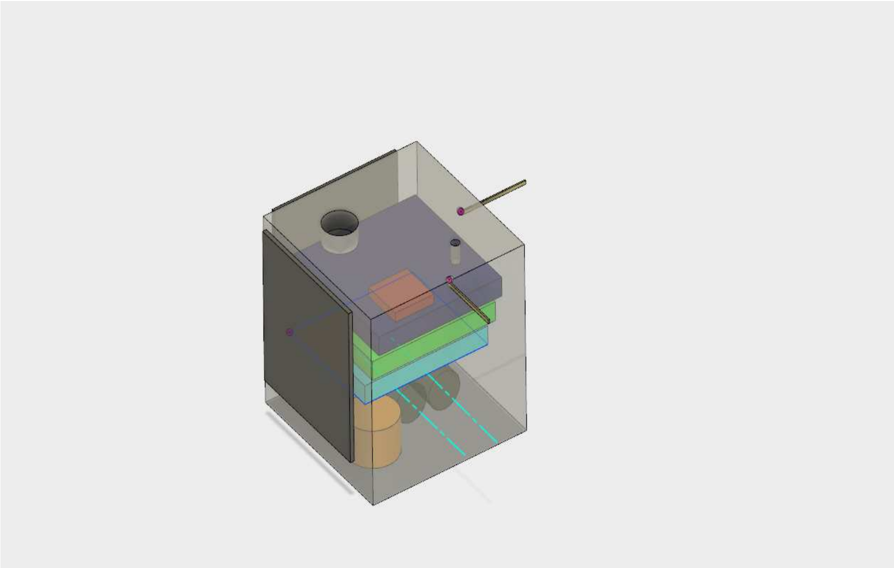


Figure (2.1) Above: Image of a CAD (Computer Aid Design) of a simple CubeSat with components integrated, without launching rails

When designing your spacecraft structure there are specific size and mass requirements you have to meet to the industry standard. The CubeSats come in typically three sizes, 1U one unit, 2U two unit, and 3U three units. You can also expand these sizes to accommodate your mission and payload specifications. These satellites are uniform and have a small for factor making them chapter to launch into space and easier to assemble and manufacture at a faster rate to meet launch deadlines and get your spacecraft into orbit quickly. In 1999 Bob Twiggs from Stanford University and Jordi Puig-Suari of California Polytechnic State University (Cal Poly Technical) developed the CubeSat for high level missions at a low cost and has now turned into a booming industry for engineers and space companies all over the world.

2.1 1U (One Unit CubeSat)

The standard 1Unit CubeSat has the dimensions of (10cm x 10cm x 10cm) or (100mm x 100mm x 100mm). This is the most standard and commonly used SmallSat in the SmallSat industry. The standard mass of the 1U CubeSat is 1.3kg and cannot exceed this number. This size CubeSat will be launched into space from rocket or International Space Station out of a customized P-POD (Poly-Picosatellite Orbital Deplorer) launcher developed by Cal Poly Technical Institute, or will be launched into orbit from a launching service company like NanoRacks. The CubeSat frame is made of aluminum chassis and has launcher rails integrated onto the four edges of the CubeSat for launch purposes and to standoff from damaging the satellite when integrating hardware, subsystems, and solar panels and also testing and transportation.

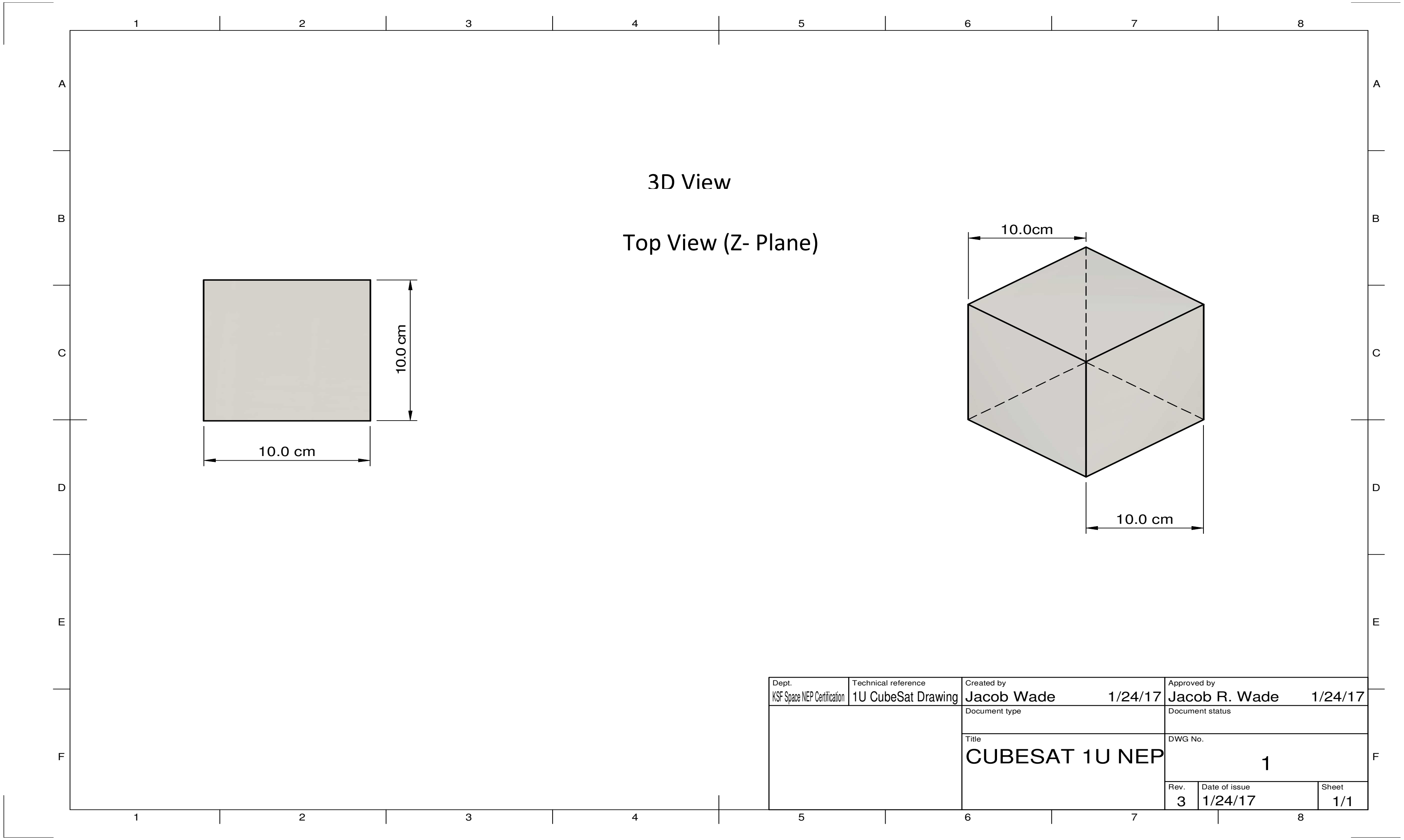


Figure (2.2) Above: This is a drawing of a 1U CubeSat labeling the dimensions of the CubeSat structure

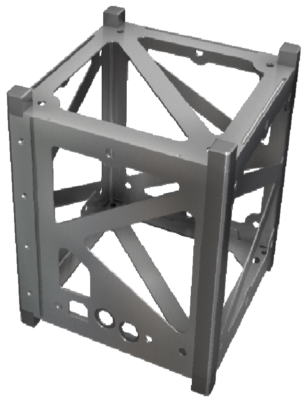


Figure (2.3) Below: 1U CubeSat chassis structure

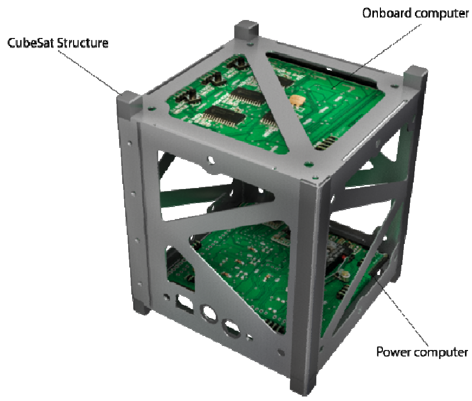


Figure (2.4) Above: Chassis 1U CubeSat frame with example PCB Components and Subsystems integrated

2.2 2 U (Two Unit CubeSat)

The 2U CubeSat is the same exact concept as the 1U but the dimensions are (100mm x 100mm x 200mm) or (10 cm x 10cm x 20cm). These CubeSats allow more space for larger payloads and propulsion systems as well. With larger components and subsystems on board there is a higher power requirement for the satellite but with the larger frame there is a larger form factor for solar panels and solar cell coverage. With more solar cells, there can also be more batteries onboard to power your spacecraft as well. The mass of the 2U should not exceed 2kg.

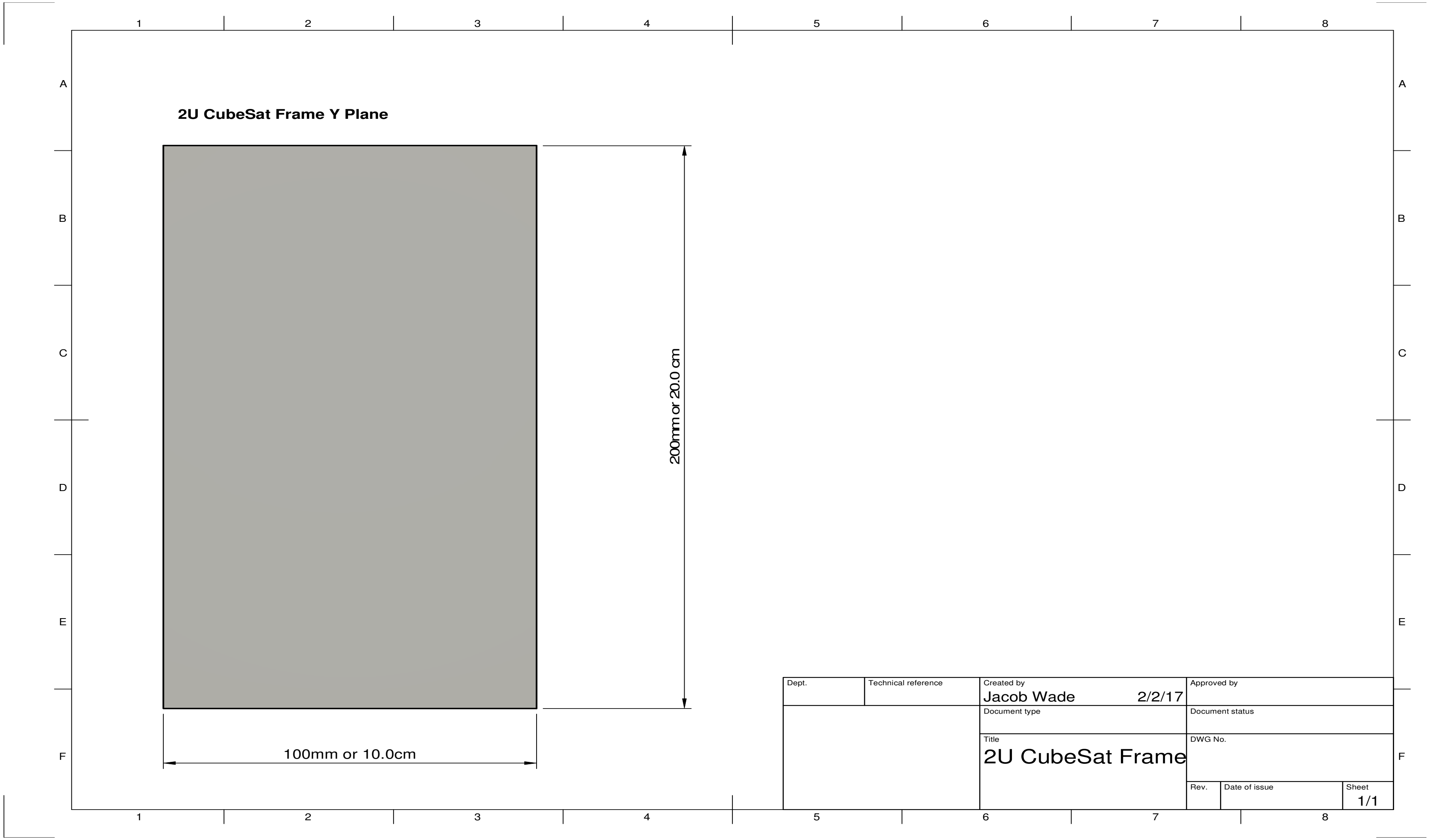


Figure (2.5) Below: Drawing of a 2U CubeSat

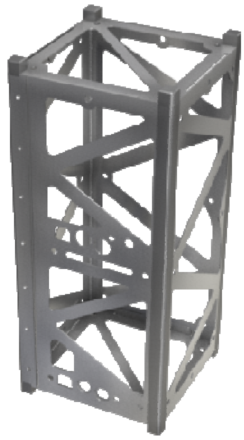


Figure (2.6) Below: Chassis 2U CubeSat aluminum frame

2.3 3 U (Three Unit CubeSat)

The 3U CubeSat is a (100mm x 100mm x 300mm) or (10cm x 10cm x 30cm) structure. Just like the 2U CubeSat the 3U has even more space for larger and multiple payloads and more advanced subsystems. For example, multiple onboard computers and power regulators and batteries, also larger radio that will provide a larger bandwidth communicating and transmitting data to the ground station. 3U CubeSats and larger units like 6U are more prone to traveling into deep space and missions to the moon and beyond HEO (High Elliptical orbit) and Earth’s gravity. With these CubeSats traveling farther into space the cost will increase because the satellite is exposed to more radiation therefore all components have to be protected and rad hardened making COTS (Commercial off the shelf) components more expansive when integrating into circuits and hardware onboard the satellite.

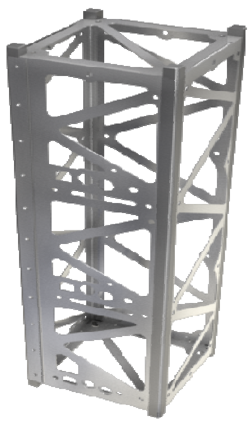


Figure (2.7) Below: 3U Chassis aluminum frame

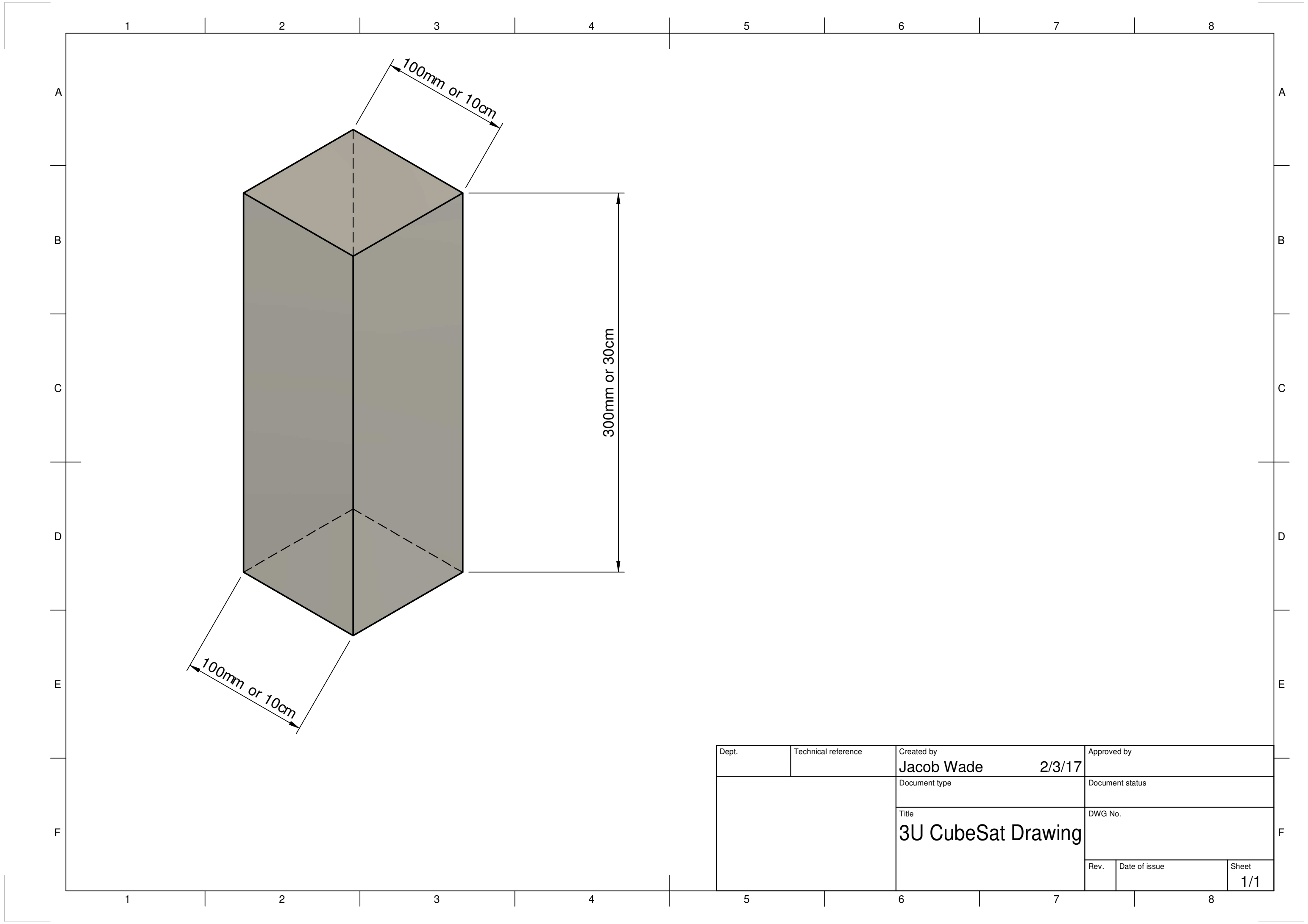


Figure (2.8) Below: Drawing of a 3U CubeSat with dimensions specified

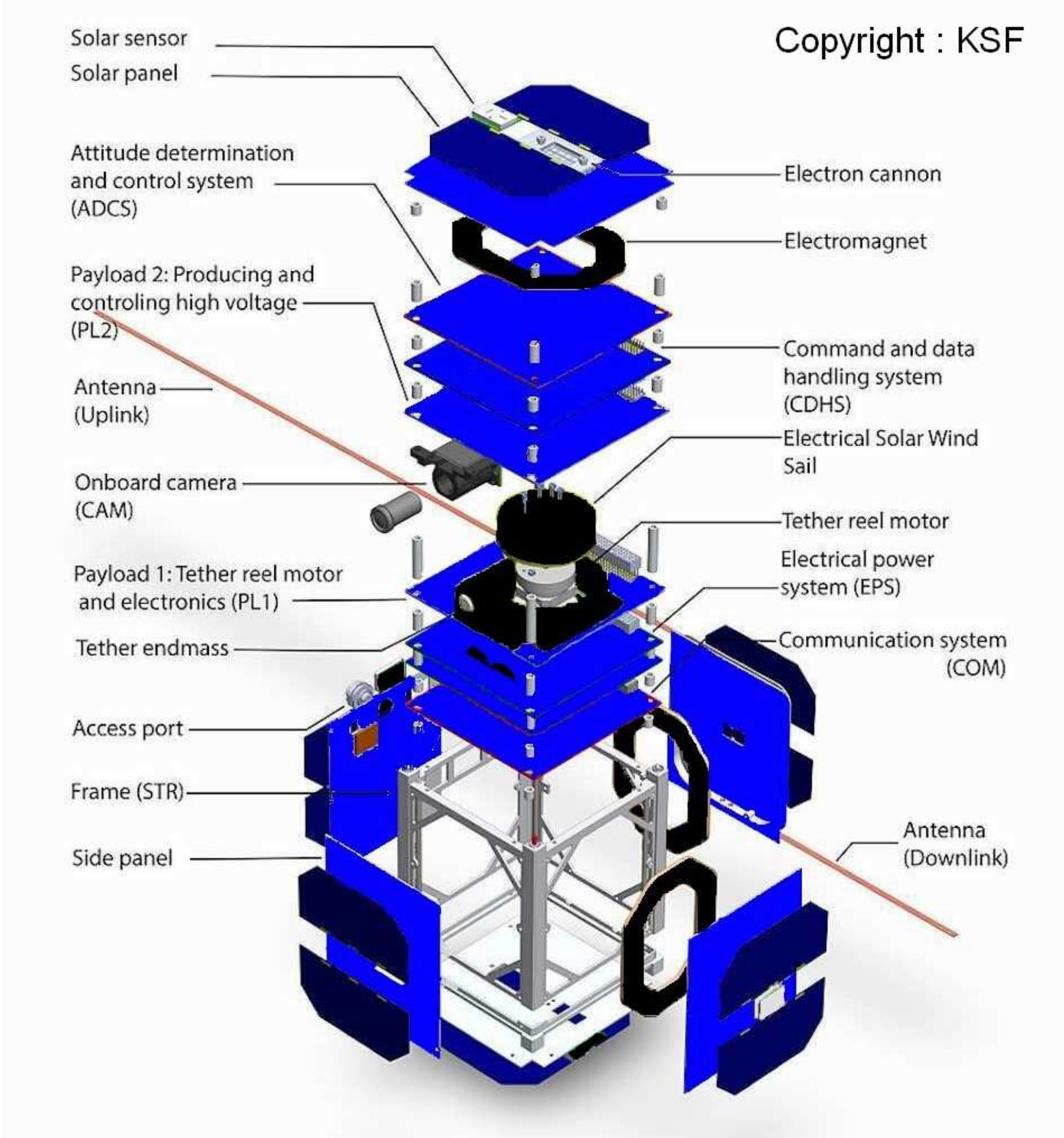


Figure (2.9) Above: 1U CubeSat CAD model (detailed) with subsystems and components labeled

2.4 CanSats

The specifications of a CanSat are the size of a 12 fluid OZ. soda can and cannot exceed a mass of 350g. CanSats are primarily designed for scientific missions into the Earth's near space atmosphere altitude, measuring air pressure, air quality, to take barometric measurements etc. These nanosatellites have COTS components (Commercial off the shelf) and are powered by solar panels or batteries. All components other than solar cells and antennas shall be internally integrated and fit inside the "can" structure. CanSats must be completely resistive to explosives, and fire because of being launched in a rocket it will reduce risk of losing your CanSat in ascent into atmosphere. The CanSat's need to be protected once deployed from the rocket and during mission because only a parachute is the primary way of descent. CanSats are the new way of collecting scientific data and measurements of Earths uppermost atmosphere and transmitting data to remote ground stations. They also have a GPS integrated for tracking ascent and descent rate and also a magnetometer to provide (X, Y, Z) position, velocity, acceleration etc. These NanoSats are durable and once you change out the batteries or re charge them and re assemble re-entry casing they are ready to fly again. These CanSats are a great way to educate young students about satellites and how they work and is a stepping stone in making a CubeSat ready for space.

Figure (2.10) Below: The CanSat structure (on right), with subsystem assembled (on left)

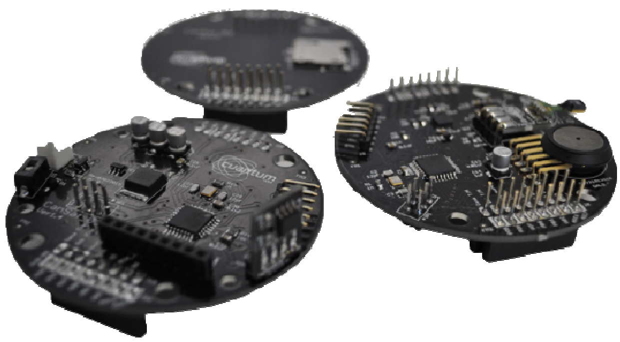


Figure (2.11) Below: CanSat Printed Circuit Boards and Integrated Circuits of payload and on-board subsystems

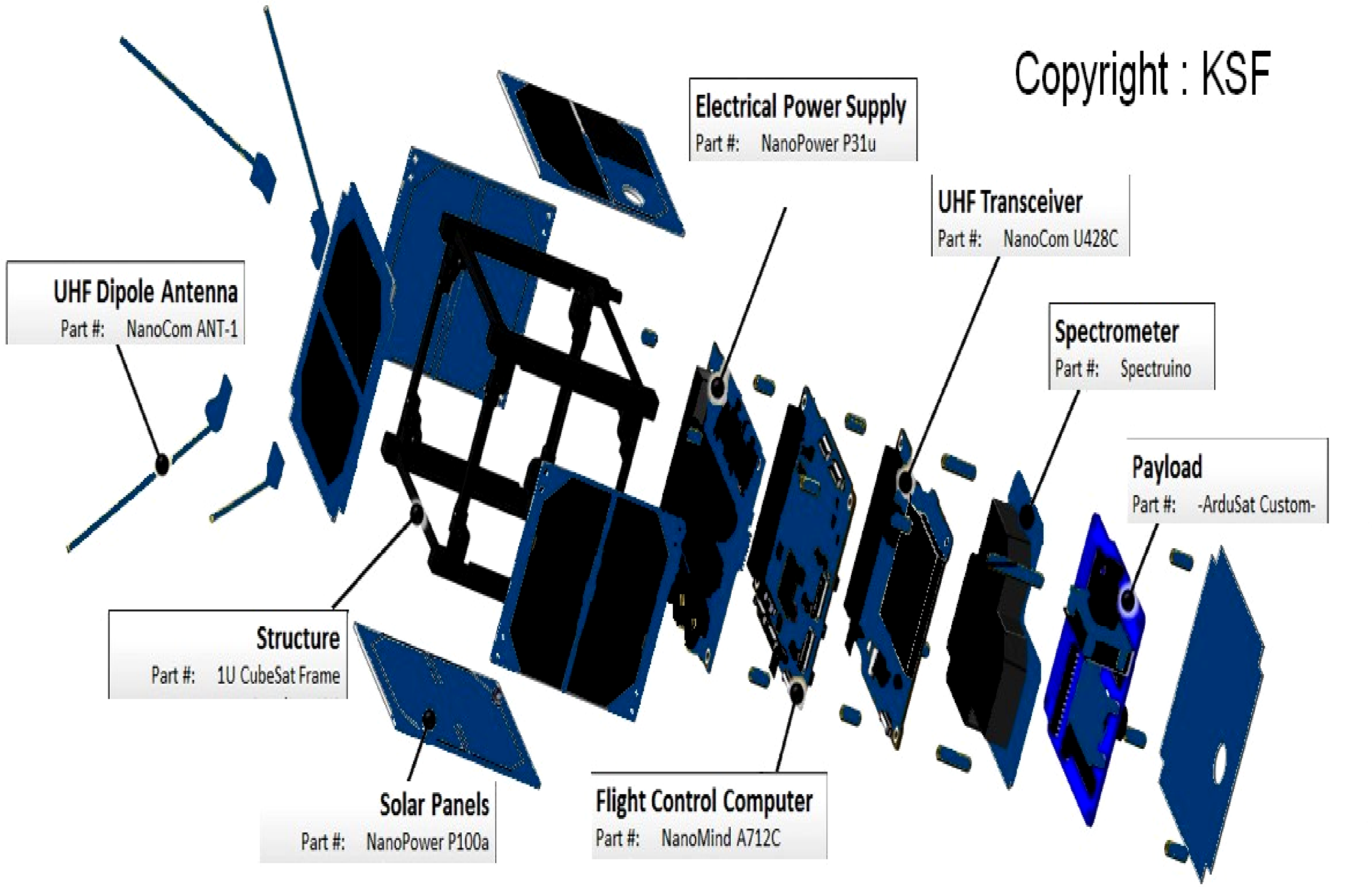


Figure (2.12) Below: Another exploded CAD model of a CubeSat

Works Cited

Pratt, Timothy, Charles W. Bostian, and Jeremy E. Allnutt. Satellite Communications. New York: Wiley, 2003. Print.

Sellers, Jerry Jon., Wiley J. Larson, William J. Astore, Anita Shute, and Dale Gay. Understanding Space: An Introduction to Astronautics. New York: McGraw-Hill, 1994. Print.

Wertz, James Richard. Space Mission Analysis and Design. El Segundo: Microcosm, 2003. Print.

Chapter 3: Component Technologies

3.1 Main computer

The main computer of a satellite is the component that manages all tasks the satellite has to do. It is responsible of maintaining the satellite operational and it monitors and surveys the other components, from the power unit to the payload, it allocates resources to other components like opening the telemetry channel to stream or receive data.

The main computer’s principal component is an integrated circuit (IC’s) in which the processing is executed. This IC’s can however belong to a different group of technologies that impact on the way the main computer is designed, its performances and its limitations.

3.1.1 Microcontrollers (Software programmable devices)

A microcontroller (uC) is the main component of an on-board computer and it communicates with more external devices likes sensors, telecommunication modules, data storage devices, etc. Every process executed in a Nano Satellite passes through the uC except when using ASIC’or FPGA’s. Microcontrollers are silicon chips with the necessary elements to execute a software program, such as memory, input/output ports and communication channels. The uC is programed using several programming languages that at first configure the chip, e.g. define how the data is going to be read, where is going to be written, how the device will communicate with peripheral devices, etc. Then the software which will be executed by the uC is developed using an Integrated Development Environment compatible with the family of the uC used.

The software or embedded systems engineer has to determine which tasks the microcontroller has to perform depending on the requirements and define the priority of such tasks to design a software that efficiently implements the application. For critical applications, such as trajectory control the microcontroller can be chosen to support a Real-Time Operating System (RTOS) and increase the reliability of the system by properly scheduling the tasks.

3.1.2 FPGA (Hardware programmable devices)

The Field Programmable Gate Arrays are silicon reprogrammable devices. They have the advantage of being flexible as a software that's is executed in a processor but without the limitations of the speed and number of cores of the processor.

The FPGA is composed of three basic elements such as Flip-Flops, logic gates (AND, OR, etc.) and wires. These elements are the building blocks of any digital circuit and the FPGA can be programmed to perform any desired task. FPGA's are the only true multitasks devices since a section of the chip can be programmed to execute a counter while another portion can be programmed to access memory and transfer data, this task will be then executed through the hardware that's be configured for that purpose all the same time.

When programming a FPGA, a binary file is created and it contains the information about which registers and logic gates will be wired up to create the logical components to perform the tasks the chip is indented to. This binary file could be created by several software programs and it is main job of the FPGA designer to described how the hardware will perform the tasks in the chip. This a totally different approach when compared with the process of writing a software program because with FPGA's the designer is describing how the operations and instructions will be executed using the reconfigurable hardware.

3.1.2.1 Tools used for configuring an FPGA

Verilog and VHDL are two hardware description languages used to describe the elements of an FPGA's and configure it. There are numerous tools used in the process of writing a hardware description using these languages, one of them is ISE 10 from Xilinx, it is an Integrated Synthesis Environments that provides the necessary elements for development and testing, it includes a simulator, a text editor, templates, a compiler, etc.

There exists another wat of configuring and programing an FPGA's using High Level Synthesis tools (HLS). For example, Lab View from NI has a library of code blocks for describing a circuit using FPGA's. In contrast with the description using the low-level description languages like VHDL or Verilog, this approach allows the designer to visualize the true nature of parallelism in FPGA's since writing a code is done line by line in a sequential fashion. The disadvantage is that this code blocks are prebuilt elements and the designer doesn't have full understanding of the final circuit which may result with not the most optimized version for such task, however this may impact only big and complex designs.

FPGA are used in specific tasks that demand lots of resources to a microcontroller in a more efficient way. For example, encoding/decoding algorithms, compression and noise filtering techniques. In difference from Application Specific Integrated Circuits (APIC's) which are chips that integrates all the necessary elements such as transistor, capacitors, resistors, etc. to perform a specific task in a single device, FPGA's can be reconfigured and are considerably less expensive than APIC's when they are used in low quantities.

3.1.3 Software and hardware programmable devices

Without a microcontroller, a computer cannot run the software that maintains the satellite alive. FPGA are used to reduce the execution time of some tasks but they don't have the capacity of a microcontroller to administrate tasks.

Currently a new kind of devices is emerging and combine elements that are software programmable and hardware programmable. This make possible a faster integration between a microcontroller and an FPGA, this way some heavy and demanding tasks can be allocated to the FPGA while the microcontroller keeps control of the resources. These new devices are currently being used in the fabrication of some main computers used in CubeSats such as the NanoMind boards from GomSpace, for example the Nano=Mind Z7000 which uses a System on Chip (SoC) Xilinx of the Zynq 7000 family which integrates an ARM processor and an FPGA. These devices are very powerful.

3.1.4 Radiation hardening devices

The Microcontrollers and FPGA's can be made more resistant for its use in space (and in other areas) by applying radiation hardening techniques to protect them from particle and electromagnetic radiation. Radiation hardened components are enhanced version of their non-hardened equivalents which have been protected with physical techniques to prevent damages on the silicon or logical techniques to correct possible malfunctions.

Rad-hardened components are however several times more expensive than their standard versions and are rarely used in Nano Satellites where the aim is to reduce costs.

3.2 Sensors

Sensors are used to convert signals from one energy form to another and are useful to measure different parameters. These devices are used in satellites for different purposes ranging from calculating the state, position and orientation of the satellite to performing scientific missions.

Usually, sensors are used as part of the payload to complete a specific mission. Depending on the mission, we can use a wide variety of sensors including optical sensors such as cameras, temperature sensors and radiation sensors.

Sensors are used also in different subsystems of the satellite including the Attitude and Orbit Control System (AOCS) and the Command and Data Handling System.

Miniaturization of technology has made possible the use of sensors in Nano-Satellites. One example of this, is the use of micro-electrical-mechanical systems (MEMS) devices.

MEMS are created by the same processes used in the semiconductor industry to produce devices with feature-sizes on the microscale.

The basic sensors in a CubeSat are:

3.3 Telemetry Module

An important aspect of a Nano Satellite is the capability to communicate and transmit scientific data to earth. The general purpose of any communication system is to maximize the data transfer rate while minimizing hardware constraints, price, and power consumption. These factors among others dictate the frequency spectrum that is appropriate for a mission.

3.3.1 Used Bands

Current small spacecraft technologies use an array of frequency bands to communicate. The majority of Nano Satellites, however, tend to use the following spectra:

- Very High Frequency (VHF) – 30 to 300 MHz
- Ultra-High Frequency (UHF) – 300 MHz to 3 GHz
- SBand–2 GHz to 4 GHz
- XBand–8 GHz to 12 GHz
- KuBand–12 GHz to 18 GHz
- KBand–18 GHz to 26.5 GHz
- KaBand-26.5 GHz to 40 GHz
- VHF/UHF

FM Radio, Television

- **S-Band**

Weather radar, surface ship radar, and some communications satellites, especially those of NASA for communication with ISS and Space Shuttle. In May 2009, Inmarsat and Solaris mobile (a joint venture between Eutelsat and Astra) were awarded each a 2×15 MHz portion of the S-band by the European Commission.

- **X-band**

Primarily used by the military. Used in radar applications including continuous-wave, pulsed, single-polarization, dual- polarization, synthetic aperture radar and phased arrays. X-band radar frequency sub-bands are used in civil, military and government institutions for weather monitoring, air traffic control, maritime vessel traffic control, defense tracking and vehicle speed detection for law enforcement.

- **Ku-band**

Used for satellite communications. In Europe, Ku-band downlink is used from 10.7 GHz to 12.75 GHz for direct broadcast satellite services, such as Astra.

- **Ka-band**

Communications satellites, uplink in either the 27.5 GHz and 31 GHz bands, and high-resolution, close-range targeting radars on military aircraft

The maximum amount of data that can be transmitted over electromagnetic waves from point A to point B depends upon the signal to noise ratio (SnR) of the system and the available bandwidth. This maximum capacity is illustrated by the Shannon-Hartley theorem and is shown in Equation 1, where C is the maximum data transfer rate in bits/second, B is the bandwidth of the channel in Hertz, and SnR is the signal to noise ratio. Equation 1 is under the assumption that the carrier frequency does not approach the data transmission rate:

$$\mathrm{C} = \mathrm{B} \cdot \log_2(1 + \mathrm{SnR}) \quad Eq\ 1$$

To increase data transmission rates there must be an increase in available bandwidth and SnR.

3.3.2 Antennas

3.3.2.1 Deployable Antennae

High gain deployable antennae are of keen interest to many small spacecraft missions.
Examples of SoA deployable antennae for small spacecraft.

Technology Name	Description	Developer	TRL Status
Deployable UHF/VHF antenna	A deployable antenna for cube satellite missions. Can deploy four monopoles antennae. Max RF power: 2 W Mass: 0.10 kg	Innovative Solutions in Space (Netherlands)	9 Flown on multiple successful missions
Deployable high gain antenna	A deployable high gain antenna for cube satellites. Max gain: 18 dBi Mass: 1.0 kg	BDS Phantomworks (USA)	6
Deployable high gain antenna	A deployable high gain antenna for cube satellites. Max gain: 15 dBi Half angle: 11deg Mass: 1.0 kg	USC's Space Engineering Research Center (SERC), (USA)	9 Launched successfully on Aeneas CubeSat

3.3.2.2 *Integrated Pointing Systems*

The current integrated pointing systems available provide a fully integrated system for a high gain antenna combined with accurate pointing units.

Examples of SoA integrated pointing systems for small spacecraft.

Technology Name	Description	Developer	TL Status
Integrated pointing high gain antenna	This technology provides a complete integrated system for a high gain antenna combined with a pointing unit accurate to 0.25 deg	Surrey Satellite Technology (UK)	9 Flown on NigeriaSat-2

3.3.2.3 *Microstrip/Patch Antennae*

There appears to be an increased use of Microstrip and Patch antennae in spacecraft communication systems. The reason for this is the microstrip and patch antenna are meant to minimize the mass and size requirements of a standard antenna while still maintaining good signal strength output. Microstrip and patch antennae are currently commercially available for a variety of frequency spectrums including the popular S-Band and X-Band.

An example of patch antennae for small spacecraft.

Technology Name	Description	Developer	TL Status
S-Band patch antenna	X-band transmitter for small satellite applications Half power angle: 70° Mass: 0.08 kg Gain: < 7 dBiC	Surrey Satellite Technology (UK)	9 Over 70 units flown

3.4 Solar Cell Overview and Physics

Solar cells are essentially a diode providing an electric current when photons interact with silicon in the p-n junction of the semiconductor. These highly energetic photons radiated from the Sun are trapped underneath the silicon layers and provide an electric current as the photons interact with the electrons in the semiconductor, making them free to move providing the voltage. Silicon is the most commonly used semiconductor because of its crystalline lattice properties making it hard for the electrons to move. When light and heat interact with silicon, there is only enough energy to move only a few electrons away from the atom, opening up a free hole for another excited electron or photon to interact with the silicon, as the freely moving electron provides a current. Since silicon is so pure with the lattice structure, other materials are doped as impurities with the silicon to have more free and excited electrons providing the electric current. Solar cells over the past decade have advanced at a rapid pace because of the material science and doping procedures adding more conductive materials to the photovoltaic cells, improving efficiency significantly.

Solar cells are very fragile and expansive yet at the same time very critical to mission success. Without solar cells the satellite would not be powered very long except for the length of the battery charge. Solar cells are the only source of power that keeps everything working onboard your satellite. On orbit the solar cells will collect solar energy when facing the sun and restore battery charge before entering shade orbit in darkness when the batteries will kick on running the satellite during this period. Solar cells are rated on efficiency and how much power output will be provided to the system in space. With commercial sized satellites used by communication companies and the military there are massive solar panels with hundreds of solar cells connected to provide the necessary power. With CubeSats and Nanosatellite's there isn't enough space to have these massive solar panels, therefore it is most important to integrate the most efficient solar cells onboard in the smallest packing form factor available. In order to achieve these high efficiency numbers, you will need to purchase from solar cell companies like SolAero and Azure Space etc.

Solar cells are becoming more and more advanced just as the CubeSats and Nanosatellites are revolutionizing the satellite industry. Majority of the solar cells available for purchase are Triple Junction solar cells (TJ). They are also integrating more materials like Gallium Arcsine on Geranium substrates causing there to be more efficient layers (junctions) that will improve solar cells efficiency. The more junctions there are the more efficient your solar cells become. Also, solar cells are integrated with a back biased diode to prevent any reverse flowing current into the other cells wired in a string together to prevent damage. In order to know how much power that needs to be provided for your satellite you must know how much the EPS and Payload and other subsystems specific power requirements are so that they are sufficiently powered throughout your mission's lifetime. The solar cells need to be tested on and IV curve (Current vs Voltage) and also a PV cure (Power vs Voltage). These two curves will characterize the solar cells so you will know how much power each solar cell will output therefore giving you the ability to accommodate the spacecraft's power requirements. Once you know how much each solar cell is expect to output you can string (wire) solar cells together you can test all solar cells together and see if it will cover the power range required.

There is another factor to take into account which is the Atmospheric Model or (AM) of each solar cell. AM0 is the model of your solar cells in space because there is no atmospheric interference from the Earth. AM1 is testing the solar cell outside on a sunny day, it is best to test solar cells outside so they are receiving actual solar energy, AM1 is the closest test to AM0 you can perform on Earth because you can't test these solar cells in space. It is best to test in AM1 on a cooler, sunny day so the heat radiating off the ground does not interfere with your current and voltage readings. Multi Meters are the most fundamental way of testing solar cells, but you can also integrate a programmable load onto the cells to see the power output from different current outputs into the cell.

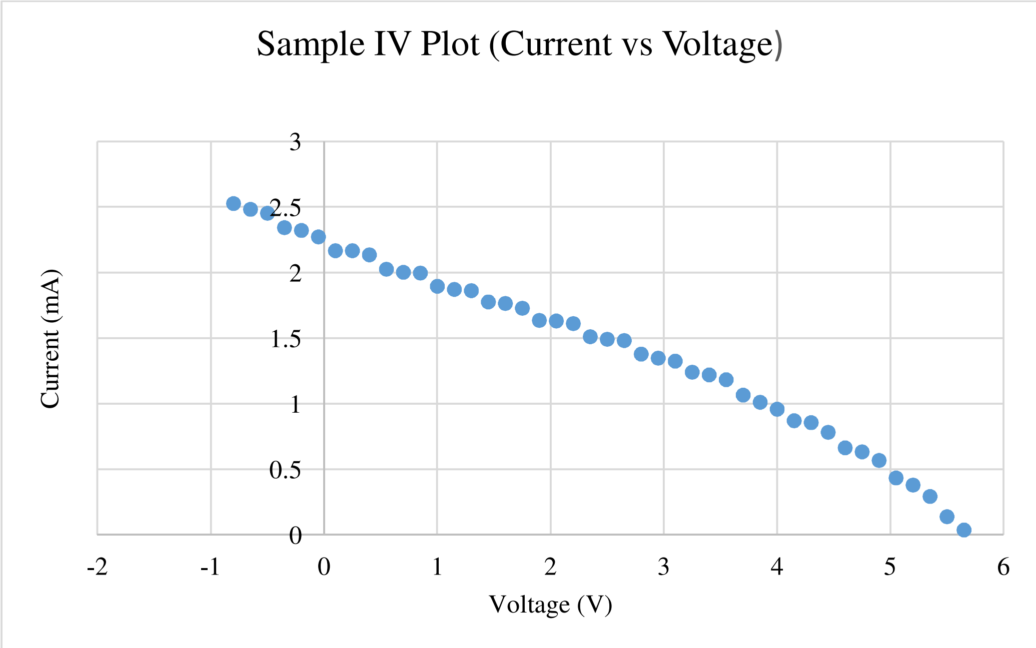


Figure (1.1) Below: Typical IV curve of a Solar Cell

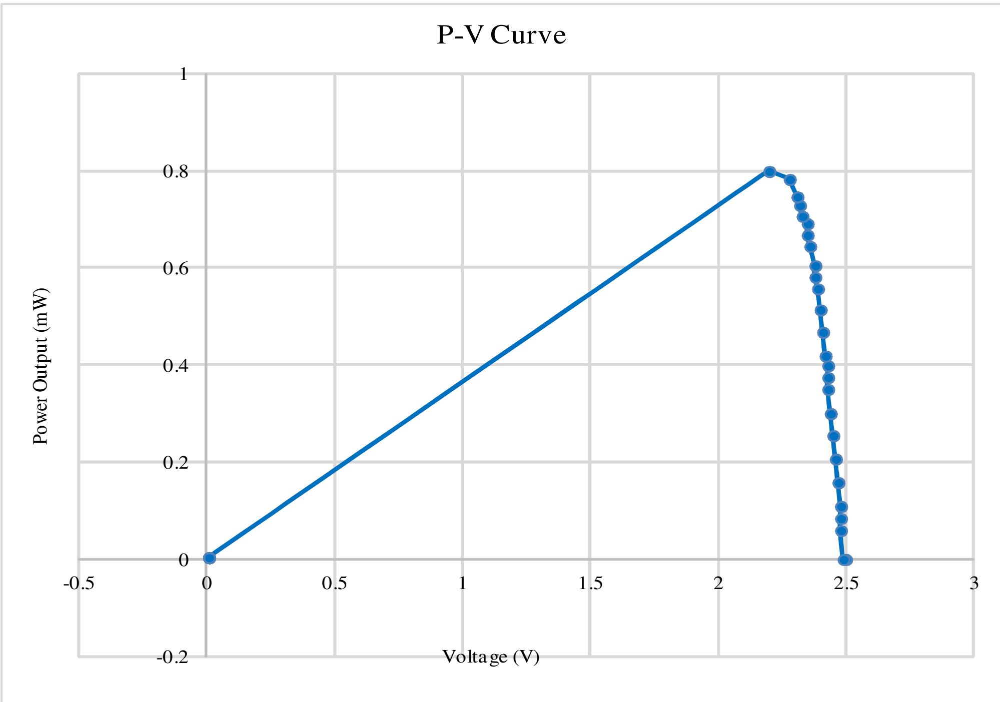


Figure (1.2) Below: Sample PV curve of a Solar Cell

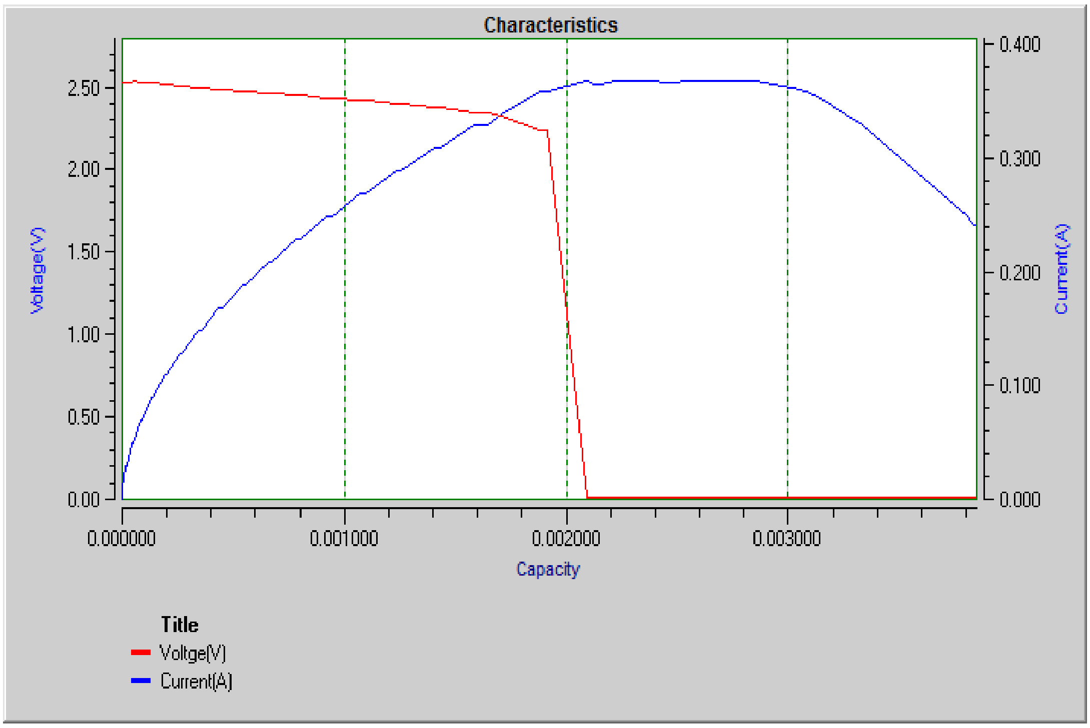


Figure (1.3) Above: Current vs Voltage plot of solar cell being characterized testing with a programmable load. (PV8500 programmable load)

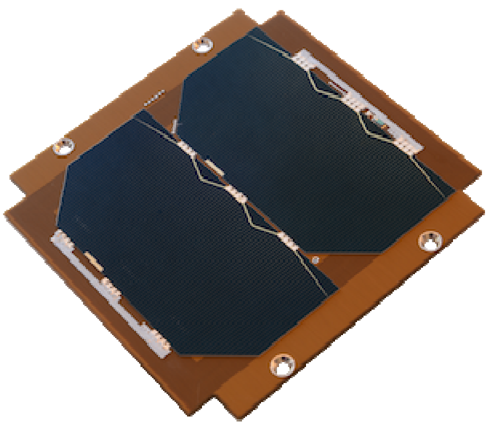


Figure (1.4) Below: CAD Model of a solar cell developed with SolidWorks

3.5 Batteries

The CubeSat power system is divided into three main areas, which include: power generation, storage, and distribution. A general layout of the power system is presented in Figure, this section is focused on the power storage.

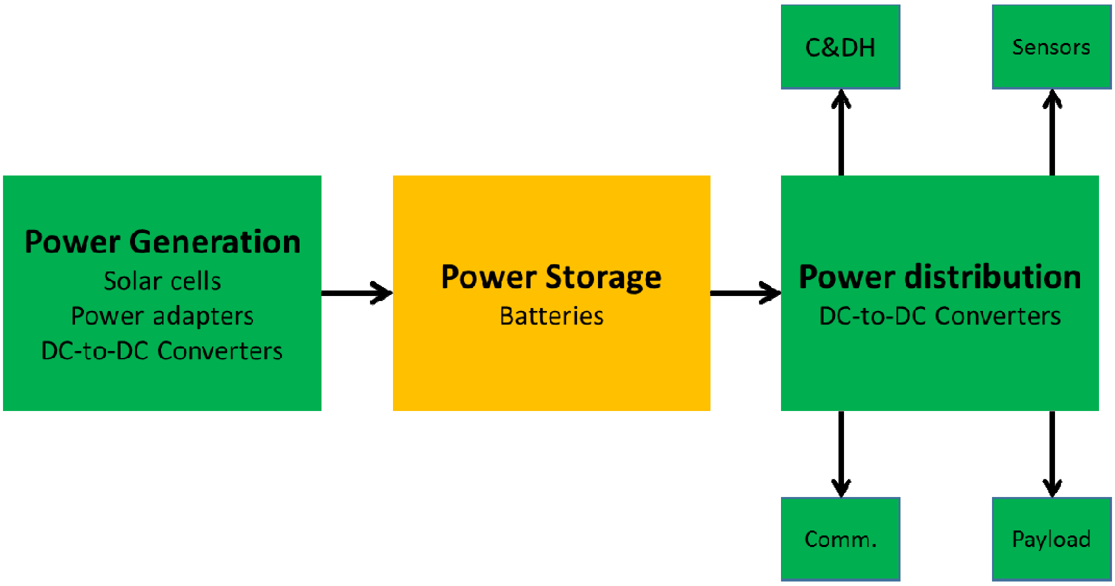
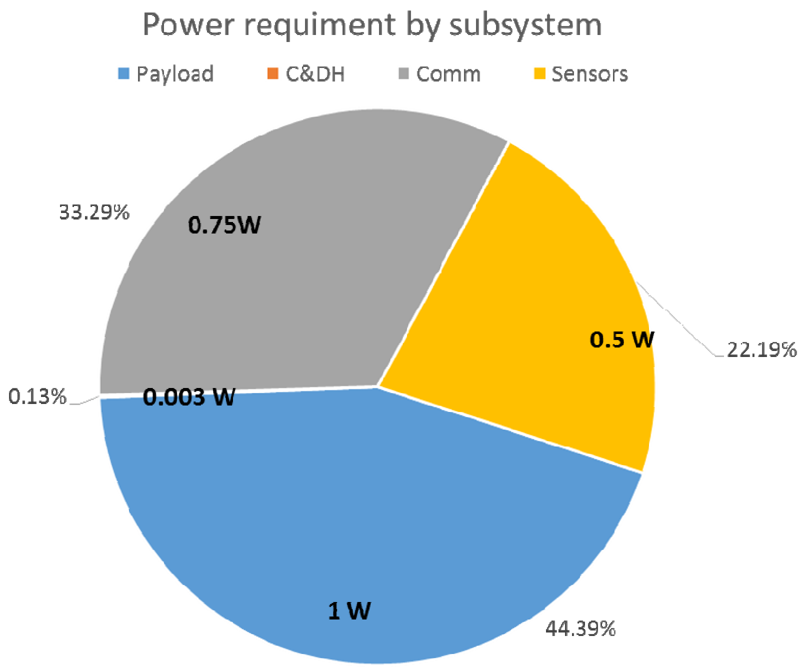


Figure (1.5) Layout of the CubeSat Power System

The power distribution element of the power system is discussed, just to give and idea of the requirements. Manufacturers of the components within aforementioned subsystems provide voltage, current, and power requirements to operate each device. The chart in Figure is an example of a distribution of power based on the requirements of other subsystems, such as the C&DH (Command and Data Handling), communications, sensors, and payload (or experiment).



For the example, the maximum power draw on the system is 2.25 Watts. To account for future changes and additions to other subsystems, the power was assumed to increase to 2.5 Watts. In addition, manufactures of the communication system, microcontroller, and GPS provided voltage requirements, which were 3.6 or 5 Volts.

Based on the maximum power requirements from each subsystem, approximately 56%, or 1.253 Watts, of the total CubeSat power is needed for the 5 Volt subsystems and 44%, or 1 Watts, for the 3.6 Volt subsystems. Therefore, the resulting minimum currents required for the 3.6 and 5 Volt supply lines are 0.278 and 0.251 Amps.

For the power storage is required a battery charger and rechargeable batteries (the most common batteries are lithium ion and Lithium polymer cells), generally, the battery charger is connected directly to the battery, which in turn, is connected to the system load. There are battery chargers that offer the capability to support the load and charge the batteries at the same time.

The power storage element is required to support the system load (power distribution) and recharge the batteries. The output voltage from the batteries must be matched with the input voltage required for the power distribution element.

3.6 PCB

3.6.1 Architecture

The final process in the manufacturing of Nano Satellites components is their integration in an electronic card known as Printed Circuit Board.

Printed circuit board is the surface with roads and ways to connect all the components between each one. This board is made typically of copper as the electrical conductor and fiber glass as the frame to support all the components.

The process begins with the design and simulation of the electrical circuit. There are a lot of software to design printed circuit boards, like OrCAD®, EAGLE®, Proteus Design Suite®, even freeware like Livewire®. Starting from the schematic diagram, where the pin pads connected between routes and each component, the selection of card dimensions and the position of the components depending of the external connectors, heat sinks, power supplies and grounding.

PCB boards can be single sided (with just one copper layer), double sided (two copper layers) or multi-layer, up to 7 layers connected with vias and the number of layers depends of the quantity of components for the requested device.

For the components, PCB’s can use two types of technology:

Through-hole technology, where the electronic components are mounted by leads inserted through holes on one side of the board and soldered with the copper traces on the other side. Leads may be soldered manually or by a wave soldering machine.

Surface-mount technology, where components became so much smaller than through-hole and have end caps that could be soldered directly into onto the PCB surface. Double sided and multi-layer PCB are more common with this technology than through-hole, and lends itself well to a high level of automation, reducing labor, costs and increasing production rates.

PCB’s are made, typically of laminates, copper-clad laminates, resin impregnated B-stage cloth and Copper foil.

Chapter 4: Software Development

4.1 Discrete Systems

A system that operates in a series of ordered steps is said to have discrete dynamics. Some systems are intrinsically discrete.

As an example, we will consider the case of an automatic counter that counts the arrivals and departures in an airport in order to know the current number of planes at any given moment. It could be modeled as shown in Figure 3.1. We won't discuss the way the count is made and will only focus on the Arrival-Detector process, which is called an "actor". Tis actor produces an event when a plane arrives, and also when a plane departs.

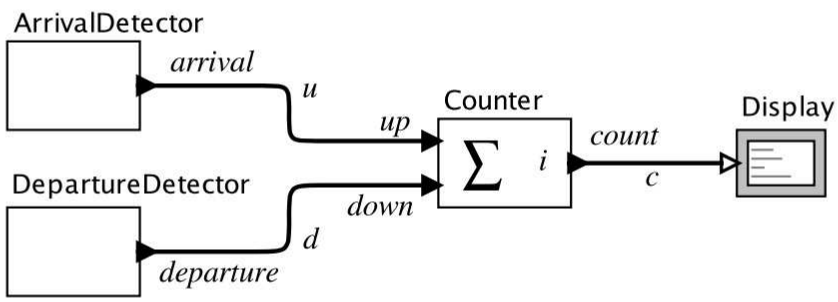


Figure 3.1: Model of a system that keeps track of the number of planes in an airport

4.2 Discrete dynamics

The actor is responsible for keeping up with the count and the events produced are used provide feedback to other components of the whole system, for example to display the current count on a screen. The behavior of this system can be described by a consequence of actions that we call reactions, which are triggered by external factors on the system environment. In this example, reactions are event triggered.

4.3 States

When talking about a state we can think about the current set of conditions that describe the system in a given time. This state how to take into account all the series of events that have led the system to that particular state, including inputs and external stimuli from the environment. Thus, the state is accumulation of past events.

The counter actor of the example has a state which the current count and that is a consequence of all past events.

The number of states the actor can take at a given time s(t) is limited to only permitted states in our example States $\subset$ Z with a limit M of planes that the airport can receive.

States = {0,1,2, ⋯, M}.

4.4 Finite-State Machines

When modeling a discrete system, a useful representation would be a model considering the discrete nature of the system. A state machine is a model that maps valuations of the inputs to valuations of the outputs, where the map may depend on its current state. A finite-state machine (FSM) is a state machine where the set States of possible states is finite.

For models with a significantly small amount of allowed states the state machine can be represented in a graphical way like in Figure 3.3. Here, each state is represented by a bubble, so for this diagram, the set of states is given by
States = {State1, State2, State3}.

At the beginning of each sequence of reactions, there is an initial state, State1, indicated in the diagram by a dangling arrow into it.

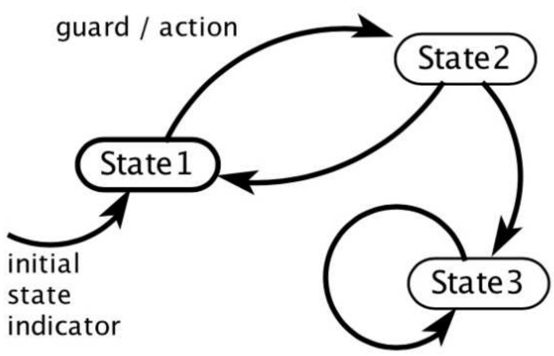


Figure 3.3: Visual notation for a finite state machine

4.4.1 Transitions

The mapping of input to output valuations is represented as arrows, called transitions, as shown in Figure 3.3. These arrows go from the actual state to the next state, after valuations are done.

When a transition starts, and ends at the same state, it's called a self-transition. This can be seen in Figure 3.3 (State3).

As seen in Figure 3.3, the arrow (transition) from State1 to State2 has the name "guard/action"; the guard tells us if the transition must be taken on a reaction, while the action tells us what outputs are to be produced on each reaction.

To be more precise, a guard is a Boolean expression that becomes true when the transition should be applied (e.g., to go from StateA to StateB). When this happens, the transition is said to be enabled.

On the other hand, an action is an assignment of outputs (values) to the reaction of the guard. It is possible that some outputs aren't present in the evaluation of the action, then we say these outputs are absent. If no output is present in the action, but they exist, we say they are implicitly absent.

4.4.2 When a Reaction Occurs

The definition itself of state machine has nothing to do with the reaction of a state, but the environment is what determines when this reaction occurs. When a reaction happens, the input values are valuated and the state machine assigns output values to this reaction, causing a state change (transition or self-transition). If a guard doesn't become true, the state machine will stay in the same state.

In a reaction, all inputs could be absent but even if this occurs, a guard could become true and then a transition happens.

When there are no inputs (absent) and no guard becomes true, the state machine is said to stutter. This kind of reaction has all inputs and outputs absent and the machine does nothing: the state is the same, no progress is made.

4.4.3 Determinacy and Receptiveness

Determinacy:

A state machine is deterministic when it has the next characteristic: for each state, there is maximum one transition enabled by each input value.

A deterministic state machine is determinate, which means that given the same inputs, the same outputs are always produced; but a determinate state machine is not always deterministic.

Receptiveness:

A receptive state machine is a machine where, for each state, there is minimum one possible transition for each input. It means that a receptive state machine is always ready to react to any input and stuttering doesn't occurs.

4.5 Hardware of embedded systems

When discussing about general-purpose computing, we can observe that the variety of instruction set architectures available is limited, with Intel x86 architecture dominating overwhelmingly. However, in the world of embedded computing there is no such dominance.

When implemented in a product, embedded processors often have a specific function. They can be used for the control of an automotive engine; process radio signals coming from space or perform recognition of faces and objects in a video stream among many other uses.

This has the advantage that processors can be more specialized and therefore more efficient performing the tasks for which they were designed. Also, there are many more benefits. For example, they can consume much less energy and therefore can be used with small batteries for long periods of time.

As a result of the wide variety of embedded applications, there are a wide variety of processors that are used. They range from very small, slow, cheap and low-power devices to high-performance and special devices.

4.5.1 Microcontrollers

A microcontroller (μC) is a small computer in a single integrated circuit consisting of a relatively simple central processing unit (CPU) combined with peripheral devices such as memories, I/O devices and timers.

The simplest microcontrollers work with 8-bit words and are suitable for applications requiring small amounts of memory and simple logic functions

They can consume extremely small amounts of energy, and often include a sleep mode that reduces energy consumption.

4.5.2 DSP Processors

It is quite common to find applications that require a bit of signal processing. An inertial measuring unit, for example, needs to read position or location information from sensors at sampling frequencies ranging from a few Hertz (Hz, or samples per second) to a few hundred Hertz. Sound applications usually require sampling rates of some kHz while video applications require smaller sampling rates but need to process large quantities of data per frame.

DSP Processors are designed to support numerically intensive signal processing applications and are capable of performing different arithmetic operations in one clock cycle.

4.5.3 Graphics Processors

Some applications like video games, video editing software and 3D modeling software require specialized processors called Graphics Processing Unit (GPU). These processors are designed to support operations with matrices which are the basis of a great majority of algorithms used in these applications.

4.6. Parallelism vs. Concurrence

Concurrency is essential for embedded systems. A computer program is said to be concurrent if different parts of the program are executed “conceptually” simultaneously. In the same way, a program is said to be parallel if different parts of the program are executed "physically" simultaneously on different hardware (such as multi-core machines, servers in a farm or different microprocessors).

Non-concurrent programs specify a sequence of instructions to execute. A programming language that expresses a calculation as a sequence of operations is called an imperative language. C is an imperative language. When C is used to write concurrent programs, we must make use of mechanisms that are not part of the language itself, usually using a thread library. A thread library uses functionalities provided not by C, but provided by the operating system and / or hardware. Java is a mainly imperative language extended with constructions that directly support threads. Thus, one can write concurrent Java programs entirely using the language characteristics.

4.7 Multitask

Threads are imperative programs that run simultaneously and share a memory space. Therefore, they can access each other's variables.

The core of a threading implementation is a scheduler that decides which thread to run when a processor is available to run a thread. The decision can be based on equity, where the principle is to give each active thread an equal opportunity to run, in the limitations of time, or in some measure of importance or priority.

The first key question is how and when the planner is invoked. A simple technique called cooperative multitasking does not interrupt a thread unless the thread itself calls a particular procedure or one of a certain set of procedures.

The main disadvantage of cooperative multitasking is that a program can run for a long time without making any operating system service calls, in which case other threads will enter a state known as starvation.

A thread can be suspended between two atomic operations to execute another thread or an interrupt service routine. This fact can make it extremely difficult to predict the behavior of the interactions between threads.

One way to prevent such disasters is through the use of a mutual exclusion lock.

A mutual exclusion lock makes the code inside the block to be executed as an atomic operation thus preventing two threads from simultaneously accessing or modifying a shared resource. The code between lock and unlock is called a critical section. At any time, only one thread can run code in such a critical section. A programmer may need to ensure that all access to a share is similarly protected by locks.

Multiprocessing programs can be very difficult to understand. In addition, it can be difficult to build trust in programs because problems in the code may not appear in the tests. A program may have the possibility of deadlock, for example, but nonetheless they run correctly for years without the lock never showing up. Programmers have to be very cautious, but reasoning about programs is difficult enough that programming errors are likely persist.

4.8 Processes

Processes are imperative programs each one with their own memory space. These programs can not refer to the variables of each one and, therefore, do not present the same difficulties as the threads. Communication between programs must be done through mechanisms provided by the operating system, microkernel or a library.

To achieve simultaneity, processes must be able to communicate. Operating systems typically provide a variety of mechanisms, including the ability to create shared memory spaces, which of course open the programmer to all possible multitasking scheduling difficulties.

One such mechanism that has the least difficulty is a file system.

A more flexible mechanism for communication between processes is the passage of messages.

Here, a process creates a piece of data, places it in a carefully controlled section of the memory being shared, and then notifies other processes that the message is ready. These other processes can block waiting for the data to be ready. Message passing requires a bit of memory to be shared, but is implemented in libraries that are presumably written by experts.

4.9 Scheduling

Real-time systems are collections of tasks in which, in addition to the ordering constraints imposed by precedencies between tasks, there are also time constraints. These restrictions refer to the execution of a task in real time. Typically, tasks have deadlines, which are time values by which the task must be completed.

A scheduler decides which task to execute next when faced with an option in running a concurrent program or set of programs.

A scheduling decision is a decision to execute a task, and has the following three parts:

- assignment: which processor should execute the task;
- order: in what order each processor must perform its tasks;
- timing: the time when each task is executed.

Each of these three decisions can be made at design time, before the program starts running, or at runtime, during program execution.

A preventive planner can make a scheduling decision during the execution of a task, assigning a new task to the same processor. That is, a task may be in the middle of execution when the programmer decides to stop that execution and start executing another task. The interruption of the first task is called preemption. A scheduler that always allows the tasks to run to completion before assigning another task to run on the same processor is called a non-preventive scheduler.

A scheduler needs some information about the structure of the program to be able to make his decisions.

The set of assumptions is called the planner's task model.

Sometimes a deadline is a real physical constraint imposed by the application, where the missing deadline is considered an error. This deadline is called a hard deadline. Scheduling with hard deadlines is called hard real-time scheduling.

Often, a deadline reflects a design decision that does not need to be strictly enforced. It is better to meet the deadline, but missing the deadline is not an error. In general, it is better not to miss the deadline by much. This case is called soft real-time scheduling.

Chapter 5: Satellite Ground Station Network and data challenges

5.1 Requirement and Strategical importance of data sharing:

While in satellite communications, ground station plays a key role to make entire operation an effective one; it is equally important to have fair data sharing policy with ground station operations. When data is transmitted to the satellite in a designated channel frequency, it is always encrypted in a language which a satellite can process for operations. The transponder in the satellite re-transmit the processed data to the ground station via sub-channel or same channel frequency, which can be decrypted at the ground station for analysis.

While running international operations in the global satellite market, where countries are performing joint operations, it is highly essential to have efficient data sharing between the service providers, end users in order to achieve the mission objectives. At this time, fair data sharing policy is required for all the connected agencies and countries having their satellite operations at the designated orbits. The ground station command network is required to be established with freeway access of data sharing for the consumers and client parties, in this case service providers and end users. With the effective access, it is strategically important to process the encrypted data in a secured way for achieving the goals of the end users, meeting their requirements. The best solution for ground station command and control network is to deliver access to not just data, but also technological advancements related to the mission for the end users, through which end users can quick and efficient decisions at low cost and higher accuracy.

5.2. Two-way communication network and policy constrains

During any satellite operations, ground station is recommended to provide two-way communication network – which here referred as between the consumers and clients. It is extremely important to have transparent and freeway communication network, especially when satellite provider and ground stations are in different countries and performing operations for other country. This scenario is common in the European Commission or BRICS or SAARC nations where multi-national agencies are involved for a single satellite operations. The major policy road block is seen while any project is involving developed and developing nations. Many times, it is seen that the bilateral agreements are restricting free access of two way communications between consumers and customers. This road block can be demolished by joint ventures or collaborative projects under various consortium, involving multi-national partners.

Another challenge is generally observed during data analysis at the ground station where English is not the first language and ground stations deliver the communication in their regional languages. This may lead to conflict in interpreting the data and communication with confusion. Hence, it is recommended to have uniform source of policy requirement for two-way open channel communication. There is a definite requirement for crystal clear legal binding in this arena of ground control network system.

5.3 Open data policy

Open data policy is a key requirement for efficient and effective decision making and provides budget relaxation for end users – especially in the case for agriculture and irrigation where farmers can utilize satellite data in a relaxed atmosphere without any major constrains of budgetary requirement or heavy finance. As per the report of Group of Earth Observation (GEO), the benefits of open data policy are in major areas but not limited to Economics of the country, Social welfare, Research and Innovations, Education and Capacity building. For example, Landsat sourced information for agriculture has been analyzed by the experts of nearly 2.0 Bn USD worldwide as per projection by ICSU.

However, there are a few important key elements which require emphasized addressing such as limits of openness, data scope and criteria...etc. These constrains can be overcome by providing relaxed ecosystem, sufficient funding, skill support and most importantly adopting mindsets for new policy. The open data policy will foster innovation in research which will ultimately benefit the society.

5.4 Global satellite traffic network

Since 1957, human race has been launching satellites into orbits and today in 2017, satellite orbits are crowded with satellites being launched every year with average of 12- 15 satellites. Also, the orbits are crowded with lots of space junk and in recent past, there have been a couple of incidences for satellite collisions with space debris. In this case, it is highly recommended to have a global satellite traffic network who monitor and track every single satellite into the orbit. This establishment unit will reduce the chances of collision between the satellites and/or between satellites and space junks. Also, it will help in maintaining clear satellite communication operations by avoiding possible conflicts and noise in the data. This can also provide information on any satellite operation which will help any developing or small island nations to acquire data during the time of disasters or any national interests.

5.5 Challenges and Opportunities:

With the advancement of technology, the challenges are also emerging on the surface for the satellite industry that includes technological, legal, policy and commercial challenges related. Major roadblock is noise disturbance in the data, which adds unnecessary time and efforts in filtering the data. However, the advancement in data fusion may overcome this challenge. While operating a satellite communication network in any collaborative project, the legal bindings must be kept clear among all the participants, data providers, end users and data processors / analyst. The policy level constrains must be addressed in order to meet cost efficient and accurate operations without any hassle. Lastly, of course budget and insurance are the most important commercial challenges which are the bottleneck of the satellite communication industry. This can easily be avoided by public-private partnership or joint ventures supported by the governments.

Exercise:

1. Explain the major challenge for data transmission at the ground station. Justify the answer with an example

2. How can we overcome difficulties related to backscattered data for agriculture in a collaborative project between developing countries, funded by a consortium?

© Copyright reserved by Dr. R Ghadawala and B R Aerospace Group, 2017

Chapter 6: Computing and Attitude Control

6.1 Introduction, Definitions and Concepts

Any spacecraft in orbit requires stabilization to increase its usefulness and effectiveness. For example, when the satellite is not stabilized it has to use omni-directional antennas so that the receivers on ground can receive the information sent by satellites irrespective of the orientation. This necessitates a high-power transmitter as only a small portion of the total power is radiated in the direction of the earth. On the other hand, if there are means to stabilize the satellite to point the antennas towards the earth, then directional antennas can be employed to increase the effective transmitter power. This reduces the actual transmitter power required compared to a satellite without stabilization. In addition, there are many other aspects which necessitate the need for satellite stabilization e.g.: solar arrays generate maximum power if they are perpendicular to the sun. and any deviation results in solar arrays generating less power. In addition, some satellites carry some scientific payloads which have to observe some point in celestial space, a particular star, the sun. a comet, etc. Moreover, when a propulsion system is also incorporated in the system, stabilization becomes a prime necessity.

Also, a satellite is subjected to a variety of disturbing torques generated by atmospheric drag, solar wind and radiation pressure, magnetic field, gravitational field and even movement of components within the satellite itself. Designers are finding ways to convert some of these disturbing torques into useful control torques. To stabilize the satellite, the attitude control becomes important due to the above requirements.

Some definitions are given next:

ATTITUDE: Orientation of a defined spacecraft body coordinate system with respect to a defined external frame.

ATTITUDE DETERMINATION: Real-Time or Post-Facto knowledge, within a given tolerance, of the spacecraft attitude Maintenance of a desired, specified attitude within a given tolerance.

ATTITUDE CONTROL: “Low Frequency” spacecraft misalignment; usually the intended topic of attitude control.

ATTITUDE ERROR: “High Frequency” spacecraft misalignment; usually unknown for the attitude control; reduced by good design or fine pointing/optical control.

A few remarks about Attitude Determination and Control Subsystem (ADCS)

● Nearly all ADCS Design and Performance can be viewed in terms of RIGID BODY dynamics

● ADCS requirements often drive overall S/C design

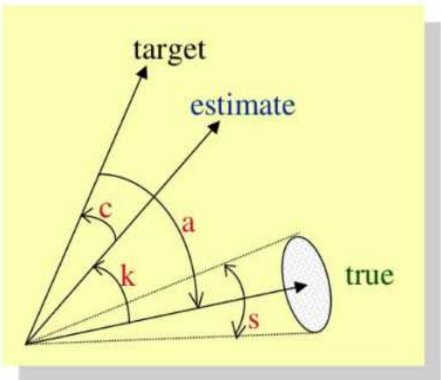
● Components of ADCS are relatively cumbersome, massive and power-consuming

● Field-of-View requirements and specific orientation are key

● Design, analysis and testing are typically the most challenging of all subsystems with the exception of payload design

● A true “systems orientation” is needed to be successful at designing and implementing an ADCS

Pointing control definitions



target	desired pointing direction
true	actual pointing direction (mean)
estimate	estimate of true (instantaneous)
a	pointing accuracy (long-term)
s	stability (peak-peak motion)
k	knowledge error
c	control error

a = pointing accuracy = attitude error
s = stability = attitude jitter

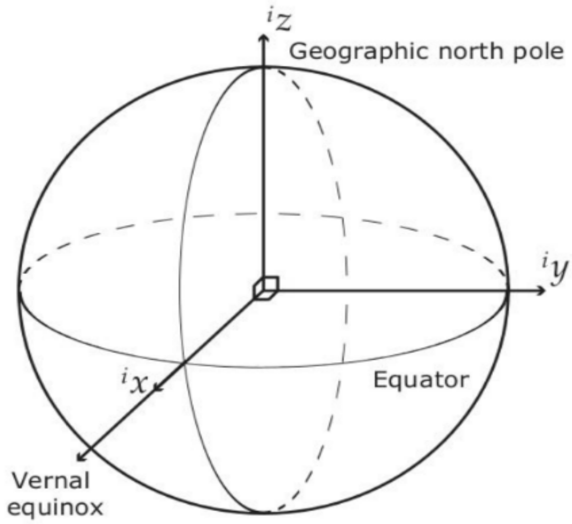
Attitude Control Implies Rotations Control, not Translations.

ADC (Attitude Determination and Control) & GNC (Guidance Navigation and Control) subsystems are often lumped together and collectively referred to ACS, ADCS or AOCS.

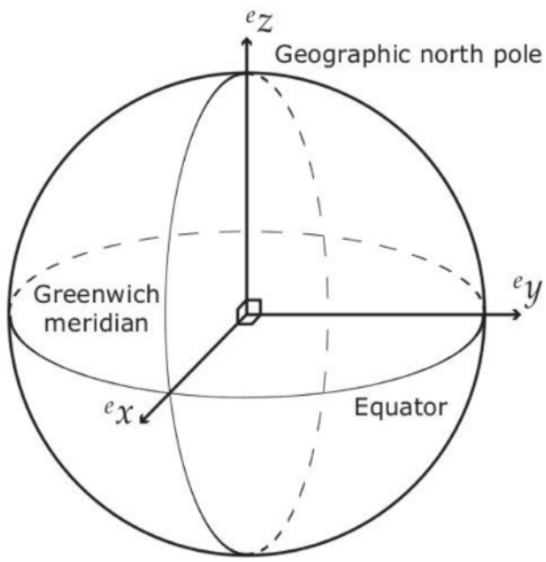
6.2 Reference frames

First of all, it is necessary to define some references for the orientation, the most commonly used reference frames will be explained next.

To define an orbit around Earth, two specific Earth related coordinate systems are given beforehand. They both have their origin in the geometrical center of Earth and are named the Earth Centered Inertial (ECI) coordinate frame and the Earth Centered Earth Fixed (ECEF) coordinate frame. These can be seen in the next figures:



(a) Earth centered inertial reference frame



(b) Earth centered Earth fixed reference frame

The ECI Coordinate Frame The term ECI stands for Earth Centered Inertial and represents a coordinate system with origin in the center of Earth, which is fixed relative to the Earth rotation. Its X-axis is parallel with the Vernal Equinox (The axis around which the Earth rotational axis is tilted relative to its orbital plane) and its Z-axis, which is parallel with the Earth rotational axis.

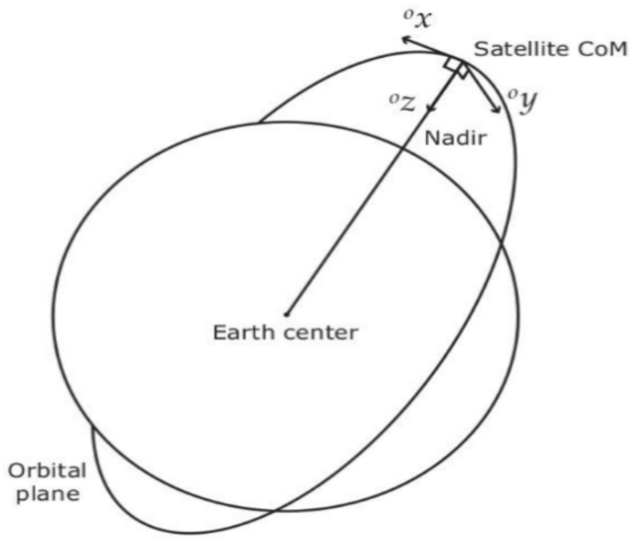
The ECEF Coordinate Frame The second coordinate frame is the Earth Centered Earth Fixed (ECEF) coordinate frame. In this frame, the X-axis is passing through the zero longitude, also known as Greenwich meridian, and has a Z-axis parallel with the rotational axis. In this way, the ECEF frame is fixed to the earth itself and rotates around with it.

Reference frames: Orbital Reference Frame

The ORF maintains its orientation relative to the Earth and follows the satellite in its orbit.

The orientation of the satellite with respect to this system of coordinates is also known as roll, pitch and yaw. The z-axis is always nadir pointing and the x-axis is parallel to the orbit plane and perpendicular to the z-axis. In a circular orbit the x-axis has the same direction as the velocity vector. Again, the y-axis is the cross product of the x- and z-axis.

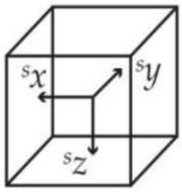
Roll is defined as the right-handed rotation about the x-axis, pitch is the rotation about the y-axis and yaw is the rotation about the z-axis.



(c) Orbit reference frame

Reference frame: Body Reference Frame

This frame is used to define orientation of ADCS hardware and attitude measurements. The x-, y- and z-axes are chosen to be parallel to the satellite frame structure, the exact definition of each axis can change according to the design of the satellite.



(d) Satellite body reference frame

To sum up the ECI and the BRF will be used for deriving the equations of motion for the satellite. The ORF is used for control purposes and to describe roll, pitch and yaw of the satellite. Lastly the ECEF is used in the calculation of the geomagnetic field and the BRF is convenient for describing position and orientation of objects like sensors and actuators.

Transformations Between Systems

After having defined reference frames for describing position and orientation of objects, it is obvious to discuss rotation of such reference frames, thus making it possible to express the orientation of the objects relative to different viewpoints.

Coordinate systems form a reference for position and angular measurement. Relationships between coordinate systems can be characterized in several methods:

- Direction Cosine Matrices.
- Euler Angle Rotation Sequence.
- Euler Parameters or quaternions.

Knowledge of the relationship between reference frames is required for attitude determination and control.

Comparison of Attitude Descriptions

Every method representation can be transformed into a different one, each of them has some advantages and disadvantages to represent an orientation or a rotation, next table gives an overview of each of them. For the computation of control techniques, the most widely used representation is the quaternion.

Method	**Euler Angles**	**Direction Cosines**	**Quaternions**
Pluses	If given ϕ, ψ, θ then a unique orientation is defined	Orientation defines a unique dir-cos matrix **R**	Computationally robust Ideal for digital control implement
Minuses	Given orient then Euler non-unique Singularity	6 constraints must be met, non-intuitive	Not Intuitive Need transforms
	Best for analytical and ACS design work		**Best for digital control implementation**

6.3. Attitude Dynamics

When talking about the control of rotations, first we have to know the differential equation, for attitude control we can reach to the following equations that describe the movements in each of the three axes:

$$\dot{H}_1 = I_1\dot{\omega}_1 = T_1 + (I_{22} - I_{33})\omega_2\omega_3$$
$$\dot{H}_2 = I_2\dot{\omega}_2 = T_2 + (I_{33} - I_{11})\omega_3\omega_1$$
$$\dot{H}_3 = I_3\dot{\omega}_3 = T_3 + (I_{11} - I_{22})\omega_1\omega_2$$

In the preceding equation, H is the angular momentum, I the matrix of inertia, ω the angular velocity and T the sum of torques applied (internal and external), I is a diagonal matrix when it is expressed in the principal reference frame.

These are known as the Euler Equations:

- No general solution exists.
- Particular solutions exist for simple torques.
- Computer simulation is usually required for the design of the ADCS.

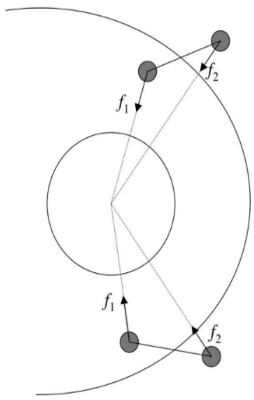
6.4 Disturbance torques

The attitude of a satellite is influenced by a number of different disturbances. These can be both internal disturbances, caused by the satellite itself, and environmental disturbances. In this section, the disturbances are also referred to as torques, since the disturbances are affecting the satellite as torques.

The internal disturbances can e.g. be caused by mass expulsion, propellant slosh and movement of mass on board the satellite. However, these disturbances are not relevant for a nanosatellite, due to the fact that very few use thrusters, propellants or moving parts on board. The dominant are external disturbances.

a) Gravity Gradient:

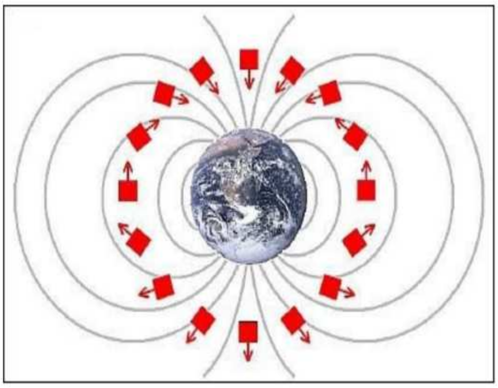
A non-symmetrical satellite in orbit about the Earth is subjected to a gravitational torque due to the Earth's non-uniform gravitational field (inverse proportional to the square of the distance to the Earth). It is important to notice that if the gravitational field of the Earth was uniform, no gravitational torque would affect the satellite.



b) Magnetic

The residual magnetic field of the satellite primarily originates from:

- Currents in the on-board electronics.
- Eddy-currents induced in the on-board electronics.
- Hysteresis effects in soft ferromagnetic materials.

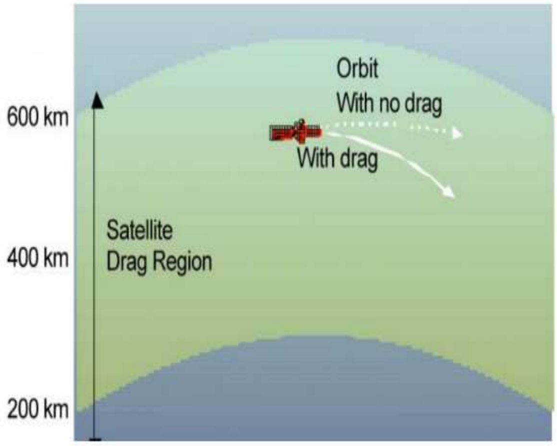


All these causes create a magnetic moment, from which the currents in the on-board electronics is the dominant source.

The magnetic residual torque T_m is caused by the interaction between the magnetic mo ment of the satellite D and the geomagnetic field B given as:

$$T_m = D \cdot B$$

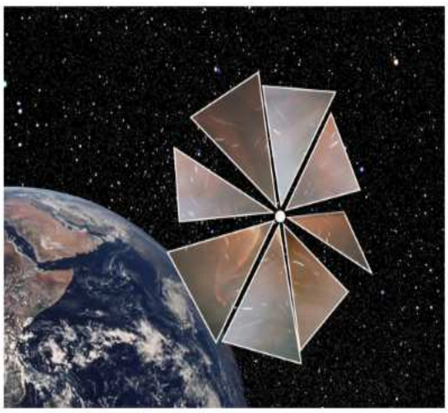
c) Aerodynamic



The aerodynamic drag disturbance originates from atmospheric molecules that collide with the surface of the satellite. These impacts are possible to model as elastic, without reflection of molecules. The drag can be modeled as a force in a point of the surface of the satellite, therefore it produces an undesired torque. Naturally, this drag force is also a perturbation for the orbit of the satellite and is stronger at lower altitudes.

The atmospheric density is not constant, but varies as a function of solar activity and whether the air is in sunlight or not. However, some models for this disturbance consider the air density to be constant.

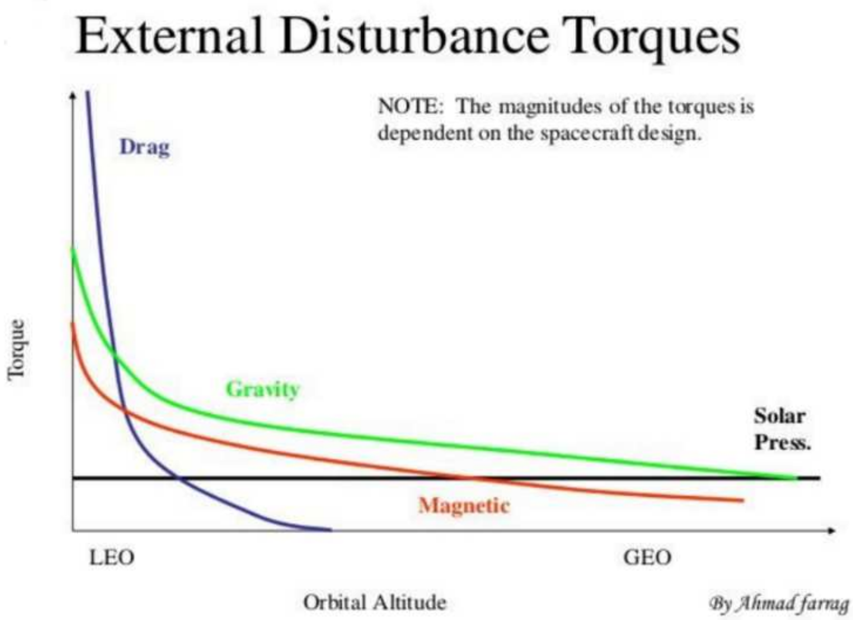
d) Solar and reflected radiation



The solar radiation is the dominating source of electromagnetic radiation. The Sun also emits particulate radiation (solar wind), which mainly consists of ionized nuclei and electrons. However, the force produced by the solar wind is negligible relative to the solar radiation. Furthermore, most of the solar wind is deflected by the Earth's magnetopause.

A sufficient method of modeling the force affecting the satellite due to solar radiation, is to assume that the incident radiation is either absorbed, reflected specularly or reflected diffuse. It is also possible to model the incident radiation as a combination of the previous mentioned cases: as a force in a point of the surface of the satellite that results in a disturbance torque.

Comparison of Disturbance torques



Usually only one of the disturbance torques is the dominant, the aerodynamic and solar radiation torque are related to the size of the structure (the surface area), for a nanosatellite with no deployables and in a low earth orbit these torques are usually negligible, even more in high orbits the aerodynamic disturbance torque is the weakest. The other torques are more significant, if a nanosatellite is very symmetric the gravity gradient torque is very small, and the magnetic disturbance torque is the dominant, the parasitic magnetic moment of the satellite can be a serious limitation in the attitude control, some measures need to be taken to reduce it in the assembly and integration stages.

6.5. Sensors and actuators

Attitude measuring sensors

As mentioned above, the attitude measuring sensors vary depending upon the mission and accuracy requirements of a satellite. Earth/horizon sensors, sun sensors, star sensors, magnetometers, and gyroscopes, are some of the sensors

Earth Sensors: Accuracy 0.1° to 1°

Earth sensors are used to scan across the earth, measuring rotation angles to define the spacecraft's altitude relative to the earth from the spacecraft's altitude. For a spacecraft in low earth orbit, the earth is the second brightest celestial object and covers up to 40% of the sky. The earth presents an extended target to a sensor compared to the Sun and stars due to their relative distances from the spacecraft. Earth emanates infrared radiation and the IR intensity in the 15 micrometer spectra! band and is relatively constant. Hence, this radiation is measured/detected using a sensor or bolometer or other type.

Sun Sensors (Accuracy: coarse 1 to 10° & fine: 0.15° to 1°)

The Sun subtends an angular radius of 0.267 degree at 1 AU (the distance between the earth and the Sun) and is nearly orbit independent as the orbital altitudes of our interest are very much smaller than 1 AU and for most applications the Sun can be treated as a point source. A simple Sun sensor can be used to detect Sun reference as the Sun is relatively very bright. A sun sensor can be used for a variety of spacecraft applications. The basic element that is most often used as the sensing element is the silicon solar cell. Whenever light falls on the sensing element, it converts the solar radiation into an electrical signal.

Star Tracker (arcseg ~ 0.01°)

Star trackers sense stars. As the stars are relatively farther than Earth and the Sun, their intensity is relatively several orders of magnitude less. These sensors usually contain a sensing element or detector, tracking platform on which the sensor is mounted or the mirror to track, a sun shade, an optical system, field of view limiter and an electronic signal processing unit, which processes the output of sensor to identify the star in association with other orbital information. Thus, these sensors provide outputs to specify the location and the visual magnitude of the star.

Magnetometer (Accuracy 2° to 5°)

Magnetometers measure the magnetic field along its input axes. Magnetometers can obtain both the direction and the magnitude of the magnetic field. Magnetometers however, are not accurate inertial attitude sensors. This is because the earth's magnetic field is not completely known and the models used to predict the magnetic field direction and magnitude at the spacecraft's position are subject to relatively substantial errors. The Earth's magnetic field strength varies with distance from the Earth as 1/r Therefore, the residual spacecraft magnetic biases will eventually dominate the total magnetic field measurement. This effect generally limits the use of magnetometers on satellites to below 1000 km.

Rate sensor (gyroscope)

Due to the power and mass limitations of CubeSats, it is not feasible to use large mechanical gyroscopes. Instead, MEMS (Microelectronic Mechanical Systems) gyroscopes must be used. MEMS gyros provide an angular rate output that must be integrated to determine the orientation of the spacecraft. MEMS gyros are small enough to fit inside a single integrated circuit package and are low power. Instead of using a spinning mass, MEMS gyros use a miniature piezoelectric oscillating mass. Motion caused by a centrifugal force due to motion disturbs the mass, and this disturbance can be correlated with angular motion.

Attitude control actuators

Attitude actuators are used to correct the attitude of a spacecraft such that it attains and stays in the desired attitude. Several attitude actuators are available that can be used depending upon the mission and payload requirements.

Thrusters (strong actuation)

Thrusters are small rockets or jets whose ON and OFF times can be controlled continuously and these are part of the Propulsion system which is dealt with in another Chapter.

Magnetic Coils (Magnetorquers/air core)

The magnetic torquer is a coil wound in ferromagnetic rod and encapsulated with a protective covering. A controlled current passing through the coil develops the desired magnetic dipole moment. These torquers are used to generate magnetic dipole moments for attitude and angular momentum control. They are also used to compensate for residual satellite biases and to counteract attitude drift due to environmental disturbance torques reacting with earth's magnetic field. To minimize volume sometimes the coil has an air core, this is mostly used in CubeSats.

Reaction Wheel (internal momentum exchange) & Momentum Wheel (momentum bias)

Momentum and reaction wheels are devices for the storage of angular momentum, which are used on spacecraft for several purposes, namely to add stability against disturbance torques, to provide a variable momentum to allow operation at one revolution per orbit for earth-oriented missions, to absorb cyclic torques, and to transfer momentum to the satellite body for the execution of slewing maneuvers.

These devices depend on the momentum of a spinning wheel. In general, a momentum wheel consists of a housing containing a flywheel, bearing assembly and electric drive motor together with electronics for driving the wheel and controlling and measuring the wheel angular rate.

6.6. Control

a) Passive Control:

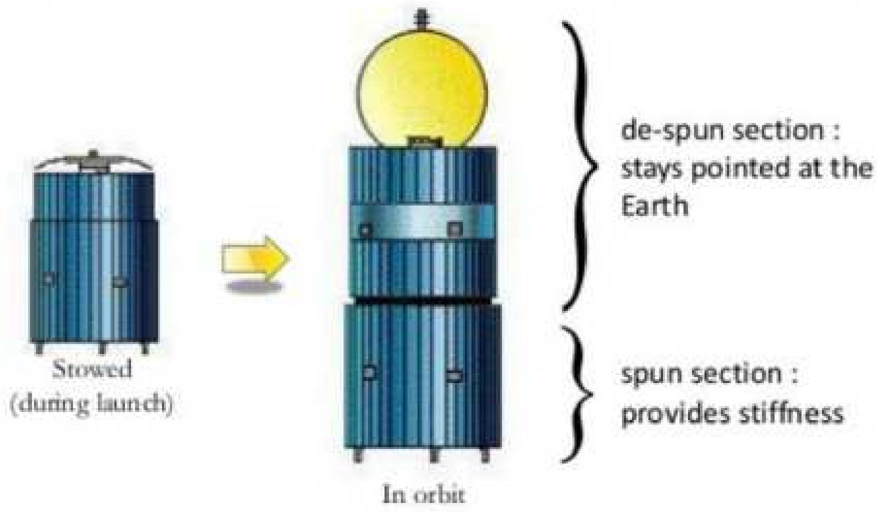
Gravity-Gradient Stabilization

In the 60's was viewed as "free" attitude control, gravity-gradient is a method of stabilizing artificial satellites or space tethers in a fixed orientation using only the orbited body's mass distribution and gravitational field. The main advantage over using active stabilization with propellants, gyroscopes or reaction wheels is the low use of power and resources.

The idea is to use the Earth's gravitational field and tidal forces to keep the spacecraft aligned in the desired orientation. In general, “GG” is not accurate enough, spacecraft can even flip over, and it is not really free, because of boom mass. However, many missions use it acceptably.

b) Active Control:

Spin / Dual Spin Stabilization



With spin stabilization, the entire spacecraft rotates around its own vertical axis, spinning like a top. This keeps the spacecraft's orientation in space under control. The advantage of spin stabilization is that it is a very simple way to keep the spacecraft pointed in a certain direction. The spinning spacecraft resists perturbing forces, which tend to be small in space, just like a gyroscope or a top. Spin-stabilized satellites must be major axis spinners: “short and fat”.

Dual spin makes the spacecraft with two parts: one spins relatively fast, fast, the other spins slowly or not at all. The major axis rule generalizes to make it possible to spin stably about the minor axis. A disadvantage to this type of stabilization is that the satellite cannot use large solar arrays to obtain power from the Sun. Thus, it requires large amounts of battery power. Another disadvantage of spin stabilization is that the instruments or antennas also must perform “despin” maneuvers so that antennas or optical instruments point at their desired targets.

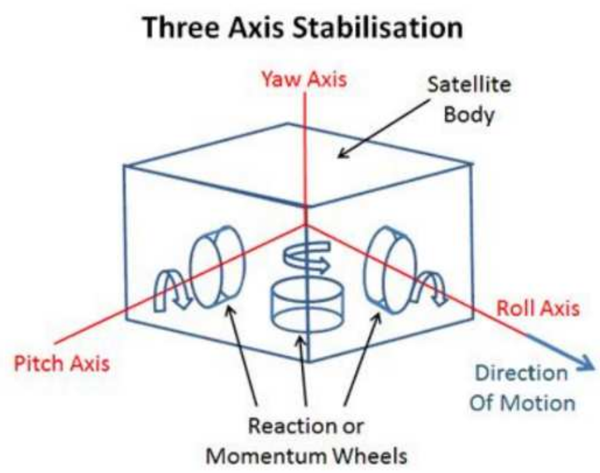
Magnetic Acquisition/Detumble System

The Magnetic Acquisition/Detumble System is an automatic despin and coarse acquisition system for three-axis stabilized spacecraft, it can be used either as a primary system, or as a backup acquisition system for a spacecraft that is inadvertently tumbling about undefinable axes at arbitrary rates. This system has been successfully used for initial despin and acquisition on several CubeSat missions. Some of the applications of this type of control are:

a) initial despin and acquisition for rotating or tumbling spacecraft:
b) autonomous backup for inertial platform control systems in the event that there is a temporary loss of control such as can be caused by a computer glitch or unexpected disturbances
c) when used in conjunction with a fixed-speed momentum wheel, the system aligns the wheel spin axis within a few degrees of the orbit normal, and reduces the pitch rate to a terminal value of twice per orbit (locked to the earth s magnetic field)

The hardware consists of a three-axis magnetometer, three magnetic torquers and a control electronics assembly. The Magnetic Acquisition/Despin System automatically reduces the rate of change of the earth's magnetic field with respect to the spacecraft in all three axes by reducing the spacecraft spin rate. This is accomplished by utilizing a three-axis magnetometer to measure the rate of change of the earth's magnetic field and then applying the proper torque to the spacecraft through the interaction of three orthogonal magnetic torquers with the earth's magnetic field.

3-axis stabilization



With three-axis stabilization, satellites have small spinning wheels, called reaction wheels or momentum wheels that rotate so as to keep the satellite in the desired orientation in relation to the Earth and the Sun. If satellite sensors detect that the satellite is moving away from the proper orientation, the spinning wheels speed up or slow down to return the satellite to its correct position. Some spacecraft may also use small propulsion-system thrusters to continually nudge the spacecraft back and forth to keep it within a range of allowed positions. An advantage of 3-axis stabilization is that optical instruments and antennas can point at desired targets without having to perform "despin" maneuvers. The obvious disadvantage is that it is much more complex and requires many types of sensors to have an accurate attitude determination prior to the attitude control part. Usually actuators are 3 or 4 reaction wheels that take considerable space inside the limited volume of a nanosatellite. However, with the appropriate design these complexities pays-off and pointing capabilities can be enhanced significantly.

Sources & References

K. F. Jensen y K. Vinther, Attitude Determination and Control System for AAUSAT-3,

Aalborg: Aalborg University Department of Electronic Systems, 2010.

A. Heydari, «Attitude Control of Magnetic Actuated Satellites, » 2007.

E. J. Øverby, «Attitude control for the Norwegian student satellite nCube, » Trondheim,

2004.

OCW 16.851 Attitude Determination and Control (ADCS) Lecture notes https://ocw.mit.edu/courses/aeronautics-and-astronautics/16-851-satellite-engineering-fall-2003/lecture-notes/l9_acs.pdf

Spacecraft Dynamics and Control - Lecture 15: Attitude Dynamics and Control http://control.asu.edu/Classes/MAE462/462Lecture15.pdf

Course AOE 4140 Spacecraft Dynamics and Control, Introduction http://www.dept.aoe.vt.edu/~cdhall/courses/aoe4140/

Chapter 7: Hall Thrusters

7.1 Introduction

Hall thrusters are relatively simple devices consisting of a cylindrical channel with an interior anode, a magnetic circuit that generates a primarily radial magnetic field across the channel, and a cathode external to the channel. However, Hall thrusters rely on much more complicated physics than ion thrusters to produce thrust. The details of the channel structure and magnetic field shape determine the performance, efficiency, and life [1–5]. The efficiency and specific impulse of flight-model Hall thrusters are typically lower than that achievable in ion thrusters [6,7], but the thrust-to-power ratio is higher and the device requires fewer power supplies to operate. The life of Hall thrusters in terms of hours of operation is usually shorter than ion thrusters (on the order of 10,000 hours), but the throughput is usually higher than in ion thrusters, and the total impulse capability can be comparable. Hall thrusters were originally envisioned in the U.S. and Russia about 50 years ago, with the first working devices reported in America in the early 1960s. Ultimately, Hall thruster technology was developed to flight status in Russia and has only recently been developed and flown outside of that country. Information about flight Hall thrusters is given in Chapter 9.

There are two generic types of Hall thrusters described in the literature. Hall thrusters, Hall-effect thrusters (HETs), stationary plasma thrusters (SPTs), and magnetic-layer thrusters are all names for essentially the same device that is characterized using a dielectric insulating wall in the plasma channel, as illustrated in Fig. 7-1. The wall is typically manufactured from dielectric materials such as boron nitride (BN) or borosil ($BN\text{-}SiO_2$) in flight thrusters, and sometimes alumina (AL_2O_3) in laboratory thrusters. These dielectric materials have a low sputtering yield and relatively low secondary electron emission coefficients under xenon ion bombardment. In this thruster geometry, the electrically biased metallic anode is positioned at the base of the channel where the majority of the propellant gas is injected into the thruster. The remainder of the propellant gas used by the thruster is injected through the exterior hollow cathode. In the second version of this type of thruster, called a thruster with anode layer (TAL), the dielectric channel wall is replaced by a metallic conducting wall, as illustrated in Fig. 7-2. This geometry considerably shortens the electric field region in the channel where the ion acceleration occurs—hence the name "thruster with anode layer" from the Russian literature [1], associated with the narrow electric field region near the anode. However, this configuration does not change the basic ion generation or acceleration method. The channel wall, which is usually also part of the magnetic circuit, is biased negatively (usually cathode potential) to repel electrons in the ionization region and reduce electron-power losses. The defining differences between these two types of Hall thrusters have been described in the literature [3].

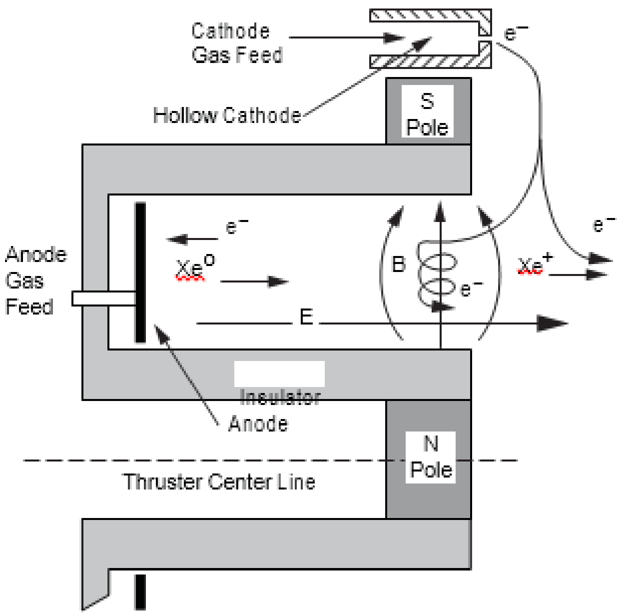


Fig. 1-1. Hall thruster cross-section schematic showing the crossed electric and magnetic fields, and the ion and electron paths

In the Hall thruster with dielectric walls illustrated in Fig. 7-1, an axial electric field is established between the anode at the base of an annular channel and the hollow-cathode plasma produced outside of the thruster channel. A transverse (radial) magnetic field prevents electrons from this cathode plasma from streaming directly to the anode. Instead, the electrons spiral along the magnetic field lines (as illustrated) and in the **E B** azimuthal direction (into the page) around the channel, and they diffuse by collisional processes and electrostatic fluctuations to the anode and channel walls. The plasma discharge generated by the electrons in the crossed electric and magnetic fields efficiently ionizes the propellant injected into the channel from the anode region. Ions from this plasma bombard and, near the channel exit, sputter erode the dielectric walls, which ultimately determines the life of the thruster. Electrons from this plasma also bombard the dielectric wall, depositing a significant amount of power in this region. The reduced axial electron mobility produced by the transverse magnetic field permits the applied discharge voltage to be distributed along the channel axis in the quasi-neutral plasma, resulting in an axial electric field in the channel that accelerates the ions to form the thrust beam. Therefore, Hall thrusters are described as electrostatic devices [1] because the ions are accelerated by the applied electric field, even though a magnetic field is critical to the process. However, since the acceleration occurs in the plasma region near the channel exit, space charge is not an issue and the ion current density and the thrust density can be considerably higher than that achievable in gridded ion thrusters. The external hollow cathode plasma is not only the source of the electrons for the discharge, but it also provides the electrons to neutralize the ion beam. The single hollow cathode in Hall thrusters serves the same function as the two cathodes in direct current (DC)-electron discharge ion thrusters that produce the plasma and neutralize the beam.

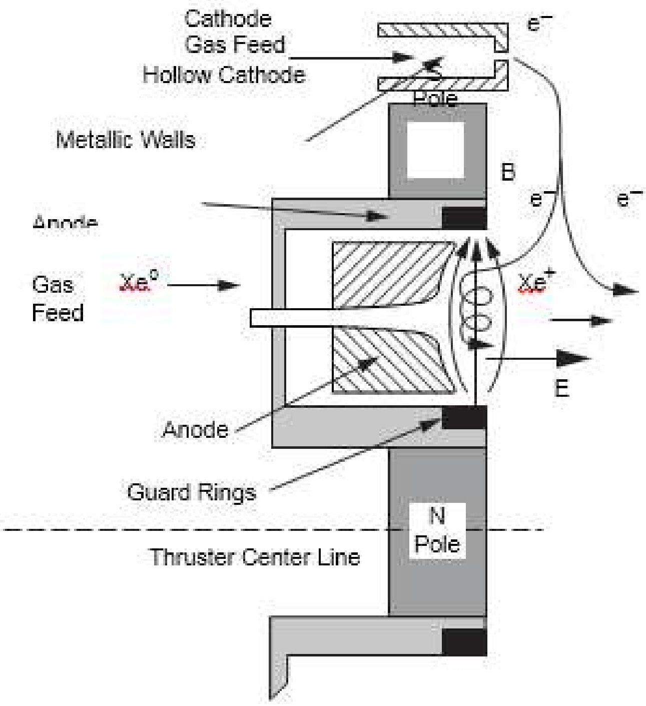


Fig. 1-2. TAL thruster cross-section schematic showing the crossed electric and magnetic fields, and the ion and electron paths

The TAL thruster with metallic walls, illustrated in Fig. 7-2, has the same functional features of the dielectric-wall Hall thruster—namely, an axial electric field is established between the anode in the annular channel and the plasma potential outside of the thruster channel. This field accelerates ions from the ionization region near the anode out of the channel. The transverse (radial) magnetic field again prevents electrons from streaming directly to the anode, and the electron motion is the same as in the dielectric-wall Hall thruster. However, the channel walls at the exit plane have metallic guard rings biased at cathode potential to reduce the electron loss along the field lines. These rings represent the major erosion source in the thruster because of ion bombardment from the plasma, and guard ring material and design often determine the thruster life. The anode typically extends close to the thruster exit and is often funnel-shaped and curved to constrain the neutral gas and plasma to the center of the channel (away from the guard rings) and to not intercept the magnetic field lines, which would cause large electron losses. However, the anode is in close proximity to the high electron-temperature region of the plasma, and electrons collected by the anode can deposit a significant amount of power. The channel width in TAL thrusters is typically twice the channel depth (including the anode shaping). The external hollow cathode plasma provides the electrons for the discharge and for neutralization the ion beam, the same as for dielectric-wall Hall thrusters.

The azimuthal drift of the electrons around the channel in the crossed electric and magnetic fields in the cylindrical thruster geometry is reminiscent of the *Hall current* in magnetron type devices, which has caused many authors to call this generically a "closed-drift" thruster [1–3]. However, King [8] correctly points out that the orientation of the fields in magnetrons (axial magnetic and radial electric) provides a restoring force to the centrifugal force felt by the electrons as they rotate about the axis, which produces the closed-drift electron motion in magnetrons.

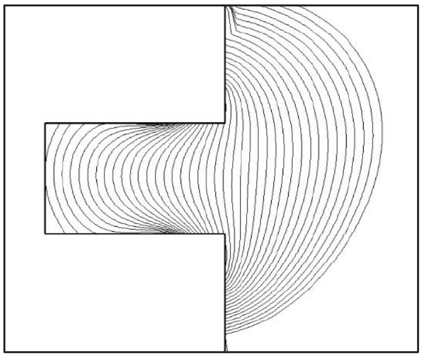
There is no corresponding restoring force associated with the different orientation of the crossed fields (radial magnetic and axial electric required to produce axial thrust) in Hall thrusters. The closed-drift behavior of the electron motion in Hall thrusters occurs only because of wall sheath electric fields and the force associated with the magnetic gradient in the radial direction in the channel, in this case, the electrons in the channel encounter an increasing magnetic field strength as they move toward the wall, which acts as a magnetic mirror to counteract the radial centrifugal force.

The radial magnetic field gradient in the channel also forms an "ion lens," which tends to deflect the ions away from the channel walls and focus the ions out of the channel into the beam. Figure 7-3 shows an example of the magnetic field lines in the NASA-173Mv Hall thruster [9] developed at the National Aeronautics and Space Administration Glenn Research Center (NASA-GRC). The curvature of the field lines in the channel approaching the exit is found to significantly improve the efficiency, especially for higher voltage, high specific impulse (Isp), Hall thrusters [9,10]. The strength of the radial magnetic field in the center along the channel [11] is shown in Fig. 1-4. The radial field peaks near the channel exit and is designed to be essentially zero at or near the anode surface.

7.2 Thruster Operating Principles and Scaling

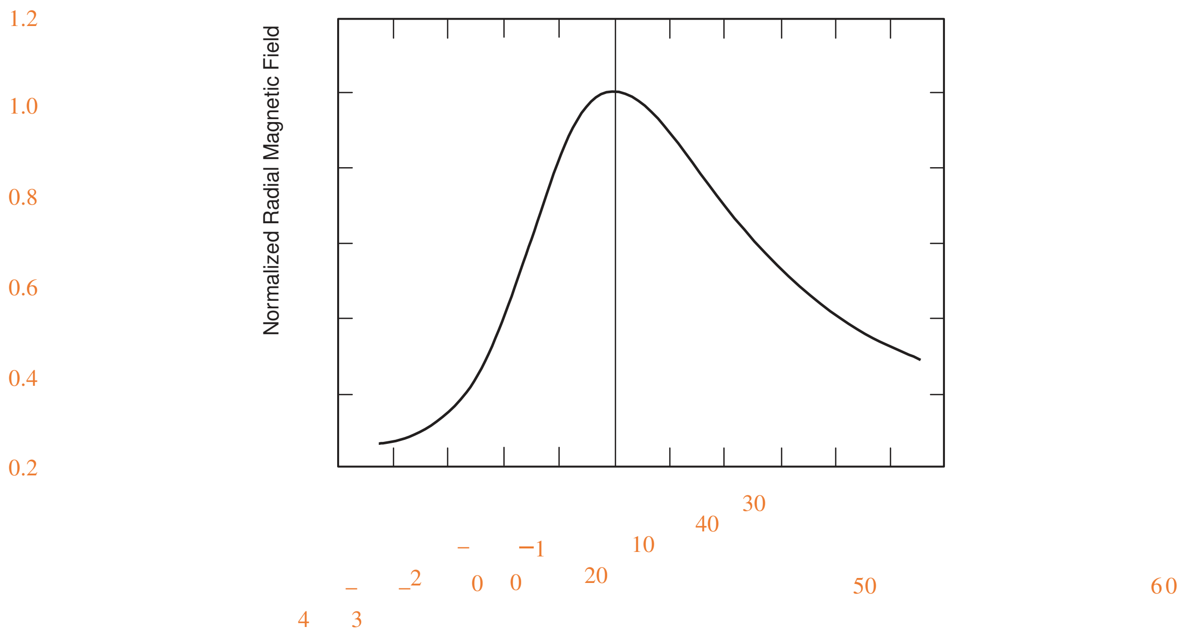
The operating principles of both types of Hall thrusters and some scaling rules for the geometries can be obtained from a simplified picture of the thruster discharge. Consider a generic Hall thruster channel, shown schematically in cross section in Fig. 7-5. The propellant gas is injected from the left through the anode region and is incident on the plasma generated in the channel. An axial scale length, L, *is* defined, over which the crossed-field discharge is magnetized, and produces a significant plasma density of width *w*, which is essentially the channel width. Ions exiting this plasma over the cylindrically

Channel



Thruster Centerline

Fig. 1-3. Magnetic field lines in the channel region of the NASA-173Mv Hall thruster (from [9])



Axial Position from Maximum *B* (mm)

Fig. 1-4. Axial variation centerline radial magnetic field normalized to the peak radial field in the NASA-173Mv Hall thruster (from [11])

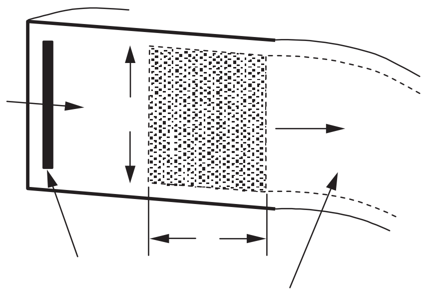
Xe^{o}

w

L

Anode

A_e



Xe^{+}

Fig. 1-5. Schematic cross section of the plasma in the Hall thruster channel

symmetric area A_e form the beam. The applied magnetic field is primarily vertical in the plasma region in this depiction.

7.2.1 Crossed-Field Structure and the Hall Current

The electrons entering the Hall thruster channel from the exterior cathode spiral around the radial magnetic field lines with a Larmor radius derived in Chapter 3 and defined by Eq. (3.3-13). The electron Larmor radius must be less than the characteristic scale length L *so* that the electrons are magnetized and their mobility to the anode is reduced. If the electron velocity is characterized by their thermal velocity, then the electron Larmor radius is

$$r_e = \frac{v_{th}}{\omega_c} = \frac{m}{eB}\sqrt{\frac{8kT_e}{\pi m}} = \frac{1}{B}\sqrt{\frac{8}{\pi}\frac{m}{e}T_{eV}} << L, \tag{1.2.1}$$

where T_{eV} is the electron temperature in eV and *L* is the magnetized plasma depth in the channel. For example, the electron Larmor radius at a temperature of 25 eV and a typical radial magnetic field strength of 150 G is 0.13 cm, which is much smaller than typical channel width and plasma length in

Hall thrusters. The electrons must also be considered magnetized, meaning that they make many orbits around a field line before a collision with a neutral or ion occurs that results in cross-field diffusion. This is normally described by stating that the square of electron Hall parameter must be large compared to unity:

$$\Omega_e^2 = \frac{\omega_c^2}{\nu^2} >> 1 \tag{1.2.2}$$

Where is the total collision frequency the effect of this criterion is clear in the expression for the transverse electron mobility in Eq. (3.6-66), where a large value for the Hall parameter significantly reduces the cross-field electron mobility.

In a similar manner, the ion Larmor radius must be much greater than the characteristic channel length so that the ions can be accelerated out of the channel by the applied electric field:

$$r_i = \frac{v_i}{\omega_c} = \frac{M}{eB}\sqrt{\frac{2eV_b}{M}} = \frac{1}{B}\sqrt{\frac{2M}{e}V_b} >> L, \tag{1.2.3}$$

Where the ion energy is approximated as the beam energy. The ion Larmor radius, for example, in the 150-G radial field and at 300 eV of energy is about 180 cm, which is much larger than the channel or plasma dimensions. These equations provide a general range for the transverse magnetic field in the thruster channel. Even if the radial magnetic field strength doubles or ion energy is half of the example given, the criteria in Eqs. (7.2-1) and (7.2-3) are still easily satisfied.

As mentioned above, the magnetic and electric field profiles are important in the thruster performance and life. The radial magnetic field typically is a maximum near the thruster exit plane, as shown in Fig. 7-4, and it is designed to fall near zero at the anode in dielectric-wall Hall thrusters [12]. Electrons from the cathode experience joule heating in the region of maximum transverse magnetic field, providing a higher localized electron temperature and ionization rate.

The reduced electron mobility and high electron temperature in the strong magnetic field region causes the axial electric field also to be maximized near the exit plane, as illustrated in Fig. 7-6. Since the neutral gas is injected from the anode region and the mass utilization is very high (nearly every neutral is ionized before reaching the channel exit), it is common to describe an "ionization region" that is located upstream of the electric field peak. Of course, the ions are accelerated directly by the electric field that peaks near the exit plane, which is sometimes called the "acceleration region." The characteristic scaling length *L* then spans these regions and is a significant fraction of the total channel depth. The ionization and acceleration regions overlap, which leads to dispersion in the ion velocity and some angular divergence in the resultant beam.

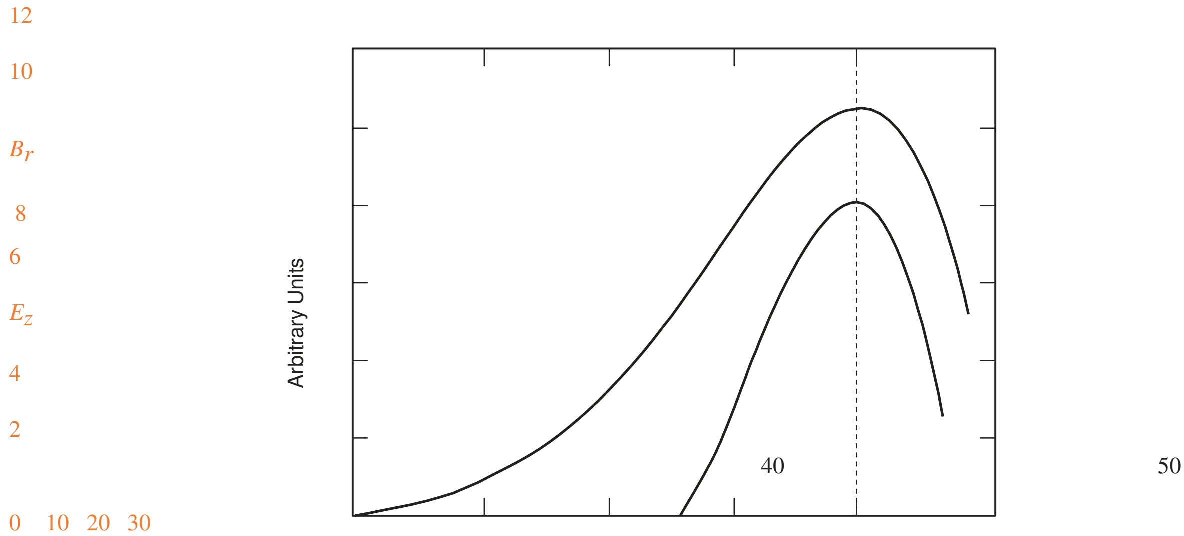
This contrasts with ion thrusters, which have a distinct ionization region in the plasma chamber and a finite acceleration region in the grids that produces nearly monoenergetic beams with low angular divergence determined by the optics and curvature of the grids.

In the crossed electric and magnetic field region of the channel, the electrons move in the azimuthal direction due to the **E** B force with a velocity given by Eq.

$$\mathbf{v} = \frac{\mathbf{E}\times\mathbf{B}}{B^2} \equiv \mathbf{v}_E , \tag{1.2.4}$$

$$v_E = \frac{\mathbf{E}\times\mathbf{B}}{B^2} \approx \frac{E_r}{B_z} \ \text{[m/s]} . \tag{1.2.5}$$

The current in the azimuthal direction, called the Hall current, is then the integral of the electron plasma density and this velocity over the characteristic thickness *L* [3,4]:



$$I_H = n_e e \left(\int_0^L v_E \, dz \right) w = n_e e \left(\int_0^L \frac{E}{B} \, dz \right) w ,$$

(1.2.6)

Where w *is* the plasma width (shown in Fig. 7-5) that essentially fills the channel. The axial electric field in the plasma channel is, approximately, the discharge voltage divided by the plasma thickness, so the Hall current is

$$I_H \approx n_e e w \frac{V_d}{B} .$$

(1.2.7)

Equation (7.2-6) shows that the Hall current increases with the applied discharge voltage and with the channel width provided that the magnetic field is unchanged. Hofer [10] showed that in Hall thrusters optimized for high efficiency, the optimal magnetic field was proportional to the discharge voltage. This implies that the Hall current is approximately constant for a given plasma density or beam current in high-efficiency Hall thrusters.

The ion current leaving the plasma to form the beam through the area A_e is approximately

$$\sqrt{I_i} = n_i e v_i A_e n_i e \frac{2eV_d}{M} 2 \, Rw , \quad (1.2\text{-}8)$$

Where R is the average radius of the plasma channel. Since the plasma is quasi- neutral (n_i n_e), even in the magnetized region, the Hall current can be expressed using Eq. (7.2-7) as

$$I_H \frac{I_i}{2\ RB} \sqrt{\frac{MV_d}{2e}}. \qquad (1.2\text{-}9)$$

Increasing the beam current in a fixed thruster design will increase the circulating Hall current for a given magnetic field and discharge voltage.

7.2.2 Ionization Length and Scaling

It is clear from the description of the Hall thruster operation above that the electrons must be magnetized to reduce their axial mobility to the anode, but the ions cannot be significantly magnetized so that the axial electric field can efficiently accelerate them to form the thrust beam. In addition, a large majority of the ions must be generated in the channel to permit acceleration by the field in that region and to produce high mass utilization efficiency [13]. This provides some simple scaling rules to be established.

The neutral gas injected from the anode region will be ionized by entering the plasma discharge in the crossed-field "ionization" region. Consider a neutral gas atom at a velocity v_n incident on plasma of a density n_e , electron temperature T_e , and thickness L. The density of the neutral gas will decrease with time due to ionization:

$$\frac{dn_n}{dt} = -n_n n_e \langle \sigma_i v_e \rangle, \qquad (1.2.11)$$

Where $\langle\sigma_i v_e\rangle$ is the ionization reaction rate coefficient for Maxwellian electrons, described in Appendix E. The flux of neutrals incident on the plasma is

$$\Gamma_n = n_n v_n , \qquad (1.2\text{-}12)$$

and the neutral velocity is $v_n = dz / dt$, where z is the axial length. equation (7.2-10) then becomes

$$\frac{d\Gamma_n}{\Gamma_n} = -\frac{n_e\langle\sigma_i v_e\rangle}{v_n} dz . \qquad (1.2.13)$$

This equation has a solution of

$$\Gamma_n(z) = \Gamma(0) e^{-z/\lambda_i} , \qquad (1.2\text{-}14)$$

Where $\Gamma(0)$ is the incident flux on the ionization region and the ionization mean free path λ_i is given by

$$\lambda_i = \frac{v_n}{n_e\langle\sigma_i v_e\rangle} . \qquad (1.2.15)$$

This expression for the ionization mean free path is different from the usual one, given in Eq. (3.6-6), that applies for the case of fast particles incident on essentially stationary particles. This is because the neutral gas atoms are moving slowly as they traverse the plasma thickness, and the fast electrons can move laterally to produce an ionization collision before the neutral leaves the region. Therefore, the ionization mean-free path depends on the neutral velocity, which determines the time the atom spends in the plasma thickness prior to a collision. The mean-free path also varies inversely with the electron density because a higher number of electrons in the slab will increase the probability of one of them encountering the neutral atom.

The percentage of the neutrals exiting the plasma of length L that are ionized is

$$\frac{\Gamma_{\text{exit}}}{\Gamma_{\text{incident}}} = 1 - e^{-L/\lambda_i} .$$

(1.2.16)

For example, to have 95% of the incident neutral flux on the plasma ionized before it leaves the plasma, Eq. (7.2-15) gives

$$L = -\lambda_i \ln(1 - .95) = 2.996\lambda_i = \frac{3v_n}{n_e \langle \sigma_i v_e \rangle},$$

(1.2.17)

or the plasma thickness must be at least three times the ionization mean-free path. Since some of the ions generated in the plasma hit the channel side walls and re-enter the plasma as neutrals instead of exiting as beam ions, the plasma thickness should significantly exceed the ionization mean-free path to obtain high mass utilization efficiency. This leads to one of the Hall thruster scaling rules:

$$\frac{\lambda_i}{L} = \text{constant} << 1 .$$

(1.2.18)

In this example, this ratio should be less than 0.33.

The actual channel's physical depth in dielectric-wall Hall thrusters is given by the sum of the magnetized plasma thickness (L) and the geometric length required to demagnetize the plasma at the anode. This is illustrated schematically in Fig. 7-6, where the channel depth is nearly twice the magnetized plasma length. The axial magnetic field gradient has been found to be critical for the thruster performance [12]. A decreasing radial magnetic field strength going toward the anode, as shown in Fig. 7-6, results in higher thruster efficiency [4,12]. At the anode, the plasma is largely un magnetized, and an anode sheath forms to maintain particle balance.

The anode sheath polarity and magnitude depend on the local magnetic field strength and direction, which affects the axial electron mobility, and on the presence of any insulating layers on the anode that affects the particle balance [14–

16]. Maintaining the local plasma near the anode close to the anode potential is important in applying the maximum amount of the discharge voltage across the plasma for the acceleration of ions. In addition, the magnetic field profile near the thruster exit strongly affects both the ability to achieve closed electron drifts in the azimuthal direction [8] and the focusing of the ions in the axial direction as they are accelerated by the electric field [9]. Optimal magnetic field design in the exit region reduces the ion bombardment of the walls and improves the ion trajectories leaving the thruster [17].

Additional information on the thruster operation can be obtained by examining the ionization criteria. Properly designed Hall thrusters tend to ionize essentially all the propellant gas incident on the plasma from the anode, so that

$$n_n n_e \langle \sigma_i v_e \rangle A_e L \approx n_n v_n A_e . \qquad (1.2\text{-}19)$$

Using Eq. (7.2-6) for the Hall current, Eq. (7.2-18) becomes

$$L = \frac{v_n V_d w}{I_H \langle \sigma_i v_e \rangle B} . \qquad (1.2.20)$$

The length of the ionization region naturally must increase with neutral velocity and can decrease with the ionization reaction rate coefficient, as seen in Eq. (7.2-16). This is important to achieve high mass utilization when propellants with a lower mass than xenon, such as krypton, are used to increase the Isp of the thruster [18,19].

Studies of optimized Hall thrusters of different sizes [20–25] have resulted in some scaling laws. A detailed comparison of the scaling laws in the literature, with experimental results from the family of empirically optimized stationary plasma thrusters (SPTs), was performed by Daren, et al. [20]. If the thruster channel inner-to-outer diameter ratio and the ionization mean-free path-to-plasma length ratio are constants, they found.

Where R is the outside radius of the channel. These scaling rules indicate that the optimum current density is essentially constant as the thruster size changes. The current density in Hall thrusters is typically in the range of 0.1 to 0.15 A/cm2.

Thus, at a given discharge voltage, the power density in a Hall thruster is also constant. Higher power densities are achieved by increasing the voltage, which has implications for the life of the thruster.

7.2.3 Potential and Current Distributions

The electrical schematic for a Hall thruster is shown in Fig. 7-7. The power supplies are normally all connected to the same reference, called the *cathode common*. The hollow cathode requires the same power supplies as an ion thruster, namely, a heater supply to raise the emitter to thermionic emission temperatures and a keeper supply for ignition and to ensure stable cathode operation at very low currents. The discharge supply is connected between the cathode common (typically also connected to the thruster body or magnetic circuit) and the anode located in the bottom of the channel. As in ion thrusters, the cathode heater is turned off once the discharge supply is turned on, and the cathode runs in a self-heating mode. The keeper is also normally used only during start-up and is turned off once the thruster is ignited. Also shown are the inner and outer magnetic field coils and their associated power supplies. Hall thrusters have been built with the cathode positioned on-axis (not shown), but this does not change the electrical schematic.

The potential distribution in a Hall thruster [26] is also illustrated in Fig. 1-7. In the upstream region of the channel where the transverse magnetic field is low, the plasma is weakly magnetized and the electron mobility is high. The plasma potential is then close to the anode potential. The plasma potential decreases toward the cathode potential near the thruster exit plane as the magnetic field increases (shown in Fig. 7-6) and limits the electron mobility. The difference between the cathode potential and the beam potential is the coupling voltage V_c, which is the voltage required to extract current from the hollow cathode. The beam voltage is then

$$V_b = V_d - V_c. \qquad (1.2\text{-}22)$$

It is common in laboratory experiments to sometimes ignore the difference in potential between the beam and ground as small (typically 10 to 20 V) and to write the beam voltage as

$$V_b = V_d - V_{cg}, \qquad (1.2\text{-}23)$$

Where V_{cg} *is* the cathode-to-ground voltage.

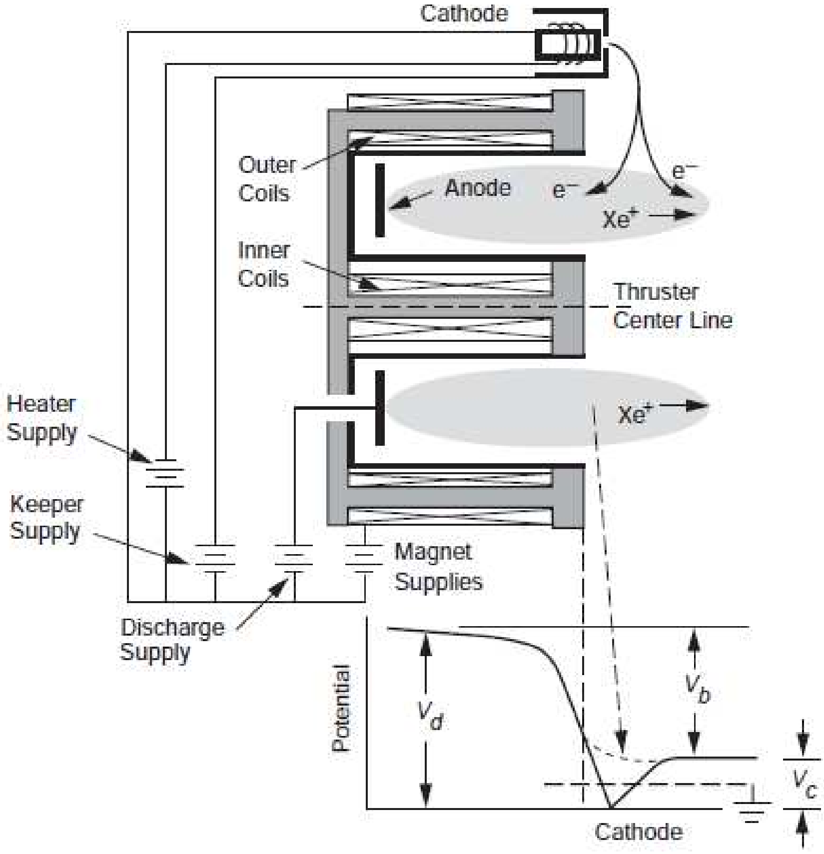


Fig. 1-7. Hall thruster electrical schematic and potential distribution

The on-axis potential, shown schematically by the dashed line in Fig. 1-7, decreases from the ionization and acceleration regions to the thrust-beam plasma potential. Ions are generated all along this potential gradient, which causes a spread in the ion energy in the beam. Since the majority of the ions are generated upstream of the exit plane (in the "ionization region"), the average velocity of the ion beam can then be expressed as

$$\langle v_b \rangle = \sqrt{\frac{2e\overline{V}_b}{M}}, \tag{1.2-24}$$

where $\overline{V}_b$ represents, in this case, the average potential across which the ions are accelerated. The actual spread in the beam energy can be significant [27,28] and must be measured by plasma diagnostics. The beam from the Hall thruster is charge neutral (equal ion and electron currents). As in ion thrusters, the thruster floats with respect to either spacecraft common in space or vacuum-chamber common on the ground. The common potential normally floats between the cathode and the beam potentials and can be controlled on a spacecraft by a resistor between the spacecraft common and the cathode common.

The actual beam energy cannot be measured directly across the power supplies because the potential difference between the beam and ground or spacecraft common is unknown and must be measured by probes or energy analyzers. The coupling voltage is typically on the order of 20 V to operate the cathode discharge properly, which usually ranges from 5% to 10% of the discharge voltage for Hall thrusters with moderate Isp. In a Hall thruster, the measured discharge current is the net current flowing through the discharge supply. The current flowing in the connection between the anode and the power supply in Fig. 7-7 is the electron and ion current arriving to the anode.

$$I_d = I_{ea} - I_{ia}. \qquad (1.2\text{-}25)$$

The ion current is typically small due to its higher mass, and so the discharge current is essentially the electron current collected by the anode. Likewise, the current flowing in the cathode leg (neglecting any keeper current) is

$$I_d = I_e + I_{ic}, \qquad (1.2\text{-}26)$$

where I_e is the emitted current and I_{ic} is the ion current flowing back to the cathode. As with the anode, the ion current to the cathode is typically small, and so the discharge current is essentially just the cathode electron emission current. Therefore, the discharge current is approximately

$$I_d = I_e = I_{ea}. \qquad (1.2\text{-}27)$$

Figure 1-8 shows a simplified picture of the currents flowing through the plasma, where the ion currents to the anode and cathode are neglected as small and the ion and electron currents to the dielectric walls are equal and are not shown. Ions are produced in the plasma by ionization events. The secondary electrons from the ionization events, I_{ei} , go to the anode, along with the primary electrons from the cathode, I_{ec} . Primary electrons either ionize neutrals or contribute energy to the plasma electrons so that the energetic electron distribution can produce the ionization. Since it is assumed that the discharge current is essentially the total electron current collected by the anode (the ion current is small), the discharge current can be written as

$$I_d = I_{ei} + I_{ec}. \qquad (1.2\text{-}28)$$

The discharge current is also essentially the electron current emitted by the cathode:

$$I_d = I_e = I_{ec} + I_{eb}. \qquad (1.2\text{-}29)$$

Using the fact that one electron and one ion are made in each ionization event such that $I_{ei} = I_{ib}$, Eq. (7.2-27) becomes

$$I_d = I_{ib} + I_{ec}. \qquad (1.2\text{-}30)$$

This relationship describes the net current crossing the exit plane, and so it is commonly stated in the literature that the discharge current is the ion beam current plus the back-streaming electron current crossing the exit plane [4,9].

Depending on the plasma conditions, it is possible for some fraction of the secondary electrons produced near the channel exit to diffuse into the beam. Equation (7.2-29) is still valid in this case because for every secondary electron that diffuses into the beam, another electron from the cathode plasma must cross the exit plane in the opposite direction to maintain the net discharge current. The discharge current is still the net ion beam current plus the backstreaming electron current across the exit plane. Finally, the ion beam current is equal to the current of electrons entering the beam: Ionization Event

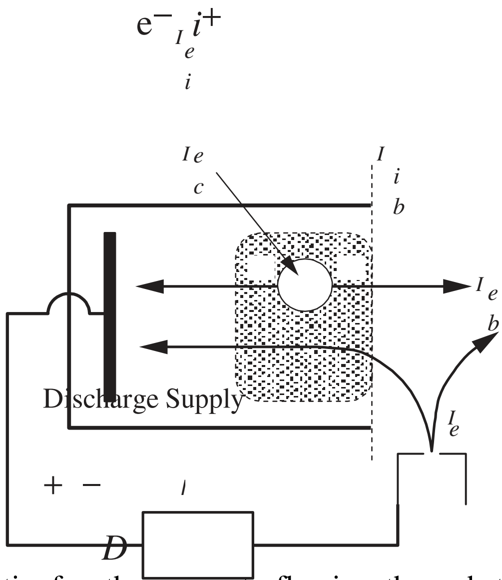


Fig. 1-8. Electrical schematic for the currents flowing through the discharge plasma and power supply

$$I_{ib} = I_{eb}. \qquad (1.2\text{-}3.1)$$

Since there is no current return path for the beam ions and electrons because the thruster floats relative to the spacecraft or the grounded vacuum system, the particles in Eq. (1.2-31) do not directly contribute to the discharge current measured in the discharge power supply.

7.3 Hall Thruster Performance Models

The efficiency of a generic electric thruster was derived in Chapter 2. Since the beam current and ion energy in Hall thrusters are not directly measured as in ion thrusters, it is useful to develop an alternative expression for the efficiency that incorporates characteristics of Hall thruster discharges. Total efficiency is always defined as the jet power, which is the thrust times the exhaust velocity, divided by the total input power:

$$\eta_T = \frac{T \ v}{P_{\text{in}}}. \qquad (1.3.1)$$

For any electric thruster, the exhaust velocity is given by Eq. (2.3-6), the Isp is given by Eq. (2.4-1), and the thrust is given by Eq. (2.3-1), which can be combined to give

$$v = \frac{\text{Isp} \cdot g}{2} = \frac{g}{2}\frac{v}{g}\frac{\dot{m}_i}{\dot{m}_p} = \frac{1}{2}\frac{T}{\dot{m}_p}. \qquad (1.3.2)$$

The total efficiency is then

$$\eta_T = \frac{T^2}{2\dot{m}_p P_{\text{in}}} \qquad (1.3.3)$$

7.3.1 Hall Thruster Efficiency

In Hall thrusters, the gas flow is split between the anode inside the discharge channel and the hollow cathode:

$$\dot{m}_p = \dot{m}_a + \dot{m}_c , \qquad (1.3\text{-}4)$$

where $\dot{m}_a$ is the anode flow rate and $\dot{m}_c$ is the cathode flow rate.

Since the cathode gas flow is injected exterior to the discharge channel ionization region and is, thereby, largely lost, the “cathode efficiency” is defined as $\eta_c = \frac{\dot{m}_a}{\dot{m}_p} = \frac{\dot{m}_a}{\dot{m}_a + \dot{m}_c}$ **(1.3.5)**

$$\eta_c = \frac{\dot{m}_a}{\dot{m}_p} = \frac{\dot{m}_a}{\dot{m}_a + \dot{m}_c} .$$

The total power into the thruster is

$$P_{in} = P_d + P_k + P_{mag} , \qquad (1.3\text{-}6)$$

where P_d is the discharge power, P_k is the cathode keeper power (normally equal to zero during operation), and P_{mag} is the power used to generate the magnetic field. The electrical utilization efficiency for the other power used in the Hall thruster is defined as

$$\eta_0 = \frac{p_d}{p_t} = \frac{p_d}{p_d + p_k + p_{mag}} \qquad (1.3.7)$$

Using Eqs. (7.3-5) and (7.3-7) in Eq. (7.3-3) gives a useful expression for the total efficiency of a Hall thruster:

$$\eta_T = \frac{1}{2} \frac{T^2}{\dot{m}_a} \frac{1}{p_d} \eta_c \eta_o \qquad (1.3.8)$$

By placing the Hall thruster on a thrust stand to directly measure the thrust, knowing the flow rates and flow split between anode and cathode, and knowing the total power into the discharge, keeper, and magnet, it is then possible to accurately calculate the total efficiency.

While Eq. (1.3-8) provides a useful expression for evaluating the efficiency, it is worthwhile to further expand this equation to examine other terms that affect the efficiency. Thrust is given from Eq.

$$T = \gamma \sqrt{\frac{2M}{e}}\, I_b \sqrt{\overline{V}_b}\,, \qquad (1.3\text{-}10)$$

where the average or effective beam voltage is used due to the spread in ion energies produced in the Hall thruster acceleration region. The fraction of the discharge current that produces beam current is

$$\eta_b = \frac{I_b}{I_d} \qquad (1.3.11)$$

Likewise, the fraction of the discharge voltage that becomes beam voltage is

$$\eta_v = \frac{v_b}{v_d} \qquad (1.3.12)$$

Inserting Eqs. (7.3-9) through (7.3-11) into Eq. (7.3-8) gives

$$\eta_T = \gamma^2 \frac{M}{e} \frac{I_d}{\dot{m}_a} \eta_b^2 \eta_v \eta_c \eta_o\,. \qquad (1.3.13)$$

Equation (7.3-12) shows that the Hall thruster efficiency is proportional to the ion mass and the discharge current, because these terms dominate the thrust production, and is inversely proportional to the anode mass flow, which dominates the mass utilization efficiency. This equation can be further simplified by realizing that

$$\frac{M}{e} I_d \eta_b = \dot{m}_i\,, \qquad (1.3.14)$$

and that the total mass utilization efficiency can be expressed as The total efficiency then becomes

$$\eta_T = \gamma^2 \eta_b \eta_v \eta_m \eta_o\,. \qquad (1.3\text{-}15)$$

This expression contains the usual gamma-squared term associated with beam divergence and multiply charged ion content and the mass utilization and electrical utilization efficiencies. However, this expression also includes the efficiencies associated with generating beam ions and imparting the discharge voltage to the beam voltage. This shows directly that Hall thruster designs that maximize beam current production and beam energy and that minimize the cathode flow produce the maximum efficiency, provided that the beam divergence and double-ion content are not adversely affected. Expressions like Eq. (7-3-15) appear in the Hall thruster literature [4,7] because they are useful in illustrating how the efficiency depends on the degree to which the thruster converts power supply inputs (such as discharge current and voltage) into the beam current and beam voltage that impart thrust. Understanding each efficiency term is critical to fully optimizing the Hall thruster performance.

The efficiency of a Hall thruster is sometimes expressed in terms of the anode efficiency:

$$\eta_a = \frac{1}{2}\frac{T^2}{\dot{m}_a P_d} = \frac{\eta_T}{\eta_o \eta_c}, \tag{1.3.16}$$

which describes the basic thruster performance without considering the effects of the cathode flow or power used to generate the magnetic field. This is usually done to separate out the cathode and magnet losses so that trends in the plasma production and acceleration mechanisms can be discerned. The anode efficiency should not be confused with the total efficiency of the thruster given by Eq. (1.3-3).

It is useful to show an example of the relative magnitude of the efficiency terms derived above. Figure 1-9 (from [10]) shows the anode efficiency that was defined in Eq. (1.3-16) and the other efficiency terms discussed above for the laboratory-model NASA-173Mv2 Hall thruster operating at 10 mg/s versus the discharge voltage. In this figure, the charge utilization efficiency is the net efficiency decrease due to multiply charged ions [10], the voltage utilization efficiency ($_v$) is the conversion of voltage into axially directed ion velocity, the current utilization efficiency ($_b$) is the fraction of ion current contained in the discharge current, and mass utilization efficiency ($_m$) is the conversion of neutral mass flux into ion mass flux. The anode efficiency increases with discharge voltage, largely because the voltage efficiency and current efficiency increase with voltage. The current utilization is always lower than the other efficiency terms, suggesting that the ultimate efficiency of Hall thrusters is dominated by the electron dynamics involved in producing the plasma and neutralizing the beam. This emphasizes the importance [9,10] of optimizing the magnetic field design to maximize the thruster efficiency.

The value of in Eq. (7.3-15) that is typically found for Hall thrusters can be evaluated using Eq. and the data in the literature. For example, a 10% double-ion content gives a thruster correction factor in Eq. (2.3-14) of = 0.973. The thrust loss due to the beam angular divergence of Hall thrusters is given by Eq. (2.3-10), (FT = cos). For both SPT-100 Hall thrusters [6] and TAL thrusters [29], a half-angle divergence of equal to about 20~deg is observed, producing FT = 0.94. The total correction factor is then = FT = 0.915 for typical Hall thruster conditions. Values for of about 0.9 have been reported.

The equivalent discharge loss for a Hall thruster can also be calculated [4,6] to provide information on how the thruster design impacts the cost of producing the beam ions. The average energy cost for producing a beam ion is the Discharge power divided by the number of beam ions minus the beam power per beam ion:

$$\varepsilon_b = \frac{I_d V_d}{I_b} - \frac{I_b V_b}{I_b} = \frac{I_d V_d}{I_b} - V_b = \frac{P_d(1-\eta_b \eta_v)}{I_b}, \tag{1.3.17}$$

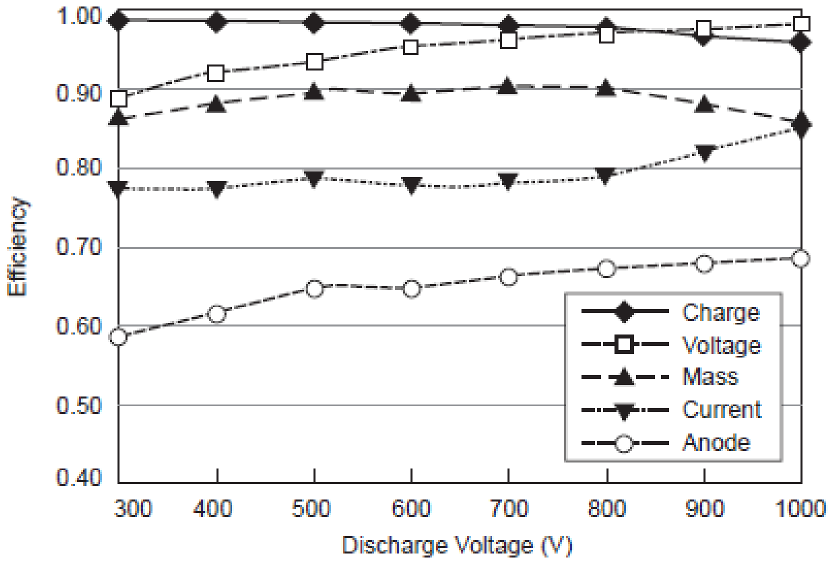


Fig. 1-9. Optimized anode efficiency and the individual efficiency terms versus discharge voltage for the NASA-173Mv2 Hall thruster operating at 10 mg/s (from [10])

Where Eqs. (1.3-10) and (1.3-11) were used. Equation (1.3-17) has the usual units for discharge loss of watts per beam-amp or electron-volts per ion. As expected, maximizing the current and voltage efficiencies minimizes the discharge loss. As an example of discharge loss in a Hall thruster, consider the SPT-100 thruster operating at the nominal 1.35-kW discharge power and 300 V. The discharge current is then 1350/300 = 4.5 A. The thruster is reported [4–6] to have values of b 0.7 and v = 0.95. The cost of producing beam ions is then:

$$\varepsilon_b = \frac{P_d\left(1-\eta_b\eta_v\right)}{I_b} = \frac{1350\left(1-0.7*0.95\right)}{0.7*4.5} = 144 \text{ [eV/ion]}.$$

7.3.2 Multiply Charged Ion Correction

In Hall thrusters operating at higher power levels (high mass flow rate and high discharge voltages >300 V), a significant number of multiply charged ions can be generated, and their effect on the performance may be noticeable. Following the analysis by Hofer [11], the performance model from the previous section can be modified to address the case of partially ionized thruster plasmas with an arbitrary number of ion species.

$$I_b = \sum_{i=1}^{N} I_i \ .$$

The total ion beam current is the sum of each ion species *i*:

The current fraction of the *i*th species is

$$f_i = \frac{I_i}{I_b}. \qquad (1.3\text{-}20)$$

Likewise, the total plasma density in the beam is the sum of the individual species densities,

$$n_b = \sum_{i=1}^{N} n_i \ , \qquad (1.3.21)$$

and the density fraction of the *i*th species is

$$\zeta_i = \frac{n_i}{n_b}.$$

$$I_b = \sum_i n_i q_i \langle v_i \rangle A_e = \sum_i n_b e \sqrt{\frac{2eV_b}{M}} \zeta_i Z_i^{3/2} \ . \qquad (1.3.22)$$

The total beam current is then, where Z_i is the charge state of each species. The mass flow rate of all the beam ion species is

$$\dot{m}_b = \frac{I_b M}{e} \sum_i \frac{f_i}{Z_i} \ . \qquad (1.3.23)$$

Using the current utilization efficiency defined in Eq. (7.3-10), the mass utilization efficiency in Eq. (1.3-14) then becomes,

$$\eta_m = \frac{\dot{m}_b}{\dot{m}_p} = \frac{\eta_b I_d M}{\dot{m}_p e} \sum_i \frac{f_i}{Z_i} . \qquad (1.3.24)$$

If the current utilization efficiency is the same for each species, then the mass utilization efficiency for arbitrary species can be written as

$$\eta_m = \eta_m^+ \sum_i \frac{f_i}{Z_i} , \qquad (1.3.25)$$

Where$_+$ is the usual ^{m} ^{m} mass utilization for a singly $^+$charged species? This is an easily implemented correction in most models if the species fractions are known. Likewise, the thrust obtained for multiple species can be generalized from Eq. (2.3-16) for Hall thrusters.

7.3.3 Dominant Power Loss Mechanisms

In preparation for examining the terms that drive the efficiency of Hall thrusters, it is useful to examine the dominant power-loss mechanisms in the thruster. Globally, the power into the thruster comes from the discharge power supply. The power out of the thruster, which is equal to the input power, is given to first order by:

$$P_d = P_b + P_w + P_a + P_R + P_{ion} , \qquad (1.3\text{-}27)$$

where P_b is the beam power given by $I_b V_b$, P_w is the power to the channel walls due to ion and electron loss, P_a is the power to the anode due to electron collection, P_R is the radiative power loss from the plasma, and P_{ion} is the power to produce the ions that hit the walls and become the beam. Additional loss terms, such as the power that electrons take into the beam, the ion power to the anode, etc., are relatively small and can usually be neglected.

In Hall thrusters with dielectric walls, the power loss due to electron and ion currents flowing along the radial magnetic field through the sheath to the channel walls (Pw) represents the most significant power loss. The current deposition and power lost to the walls can be estimated from the sheath potentials and electric fields in the plasma edge. Since the wall is insulating, the net ion and electron currents to the surface must be equal. However, ion and electron bombardment of common insulator materials, such as boron nitride, at the energies characteristic of Hall thrusters produces a significant number of secondary electrons, which reduces the sheath potential at the wall and increases the power loading.

The requirement of local net current equal to zero and particle balance for the three species gives:

$$I_{iw} = I_{ew} - \gamma I_{ew} = I_{ew}\left(1-\gamma\right), \tag{1.3.28}$$

where is the secondary electron yield from electron bombardment? Using Eq. for the Bohm current of ions to the wall, Eq. (3.7-52) for the electron current to the wall, and neglecting the secondary electron velocity, Eq. (7.3-28) can be solved for the sheath potential s , including the effect of secondary electron emission:

$$\phi_s = \frac{kT_e}{e} \ln\left[\left(1-\gamma\right)\sqrt{\frac{2M}{\pi m}}\right]. \tag{1.3.29}$$

This expression is slightly different than that found in the literature [30,31] because we have approximated e–1/2 = 0.61 0.5 for the coefficient in the expression for the Bohm current. Nevertheless, as the secondary electron yield increases, the sheath potential decreases from the classic floating potential described in Chapter 3 toward the plasma potential.

Secondary electron yields reported in the literature [30,32,33] for several materials used for the walls of Hall thrusters are shown in Fig. 7-10. In this figure, the measurements were made using a monoenergetic electron gun. Generalizing these data for incident Maxwellian electron temperatures is accomplished by integrating the yield over the Maxwellian electron energy distribution function, which results in multiplying the secondary emission scaling by the gamma function [30]. An expression for the secondary electron yield from electron bombardment of materials is then

$$\gamma = \Gamma(2+b) a T_{\mathrm{eV}}^{b}, \tag{1.3.30}$$

Where the electron temperature is in electron volts, (x) is the gamma function, and the coefficients a and b are found from fits to the data in Fig. 7-10. Values of the coefficients in Eq. (7.3-30) can be found in Table 7-1 for these materials, and the actual secondary electron yield for the Hall thruster walls is plotted versus plasma electron temperature in Fig. 7-11. It should be noted that due to reflection at the wall, the effective secondary electron yield does not go to zero for zero electron energy. This effect is accommodated by linear fits to the data

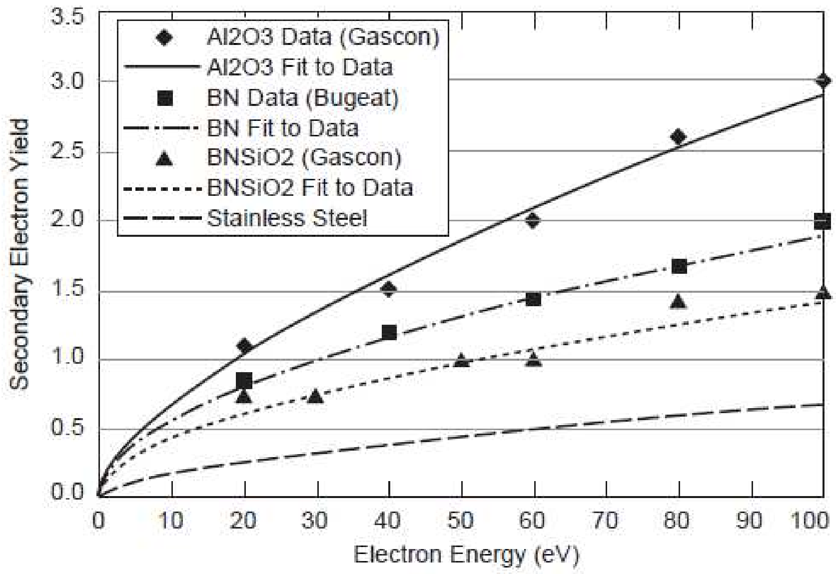
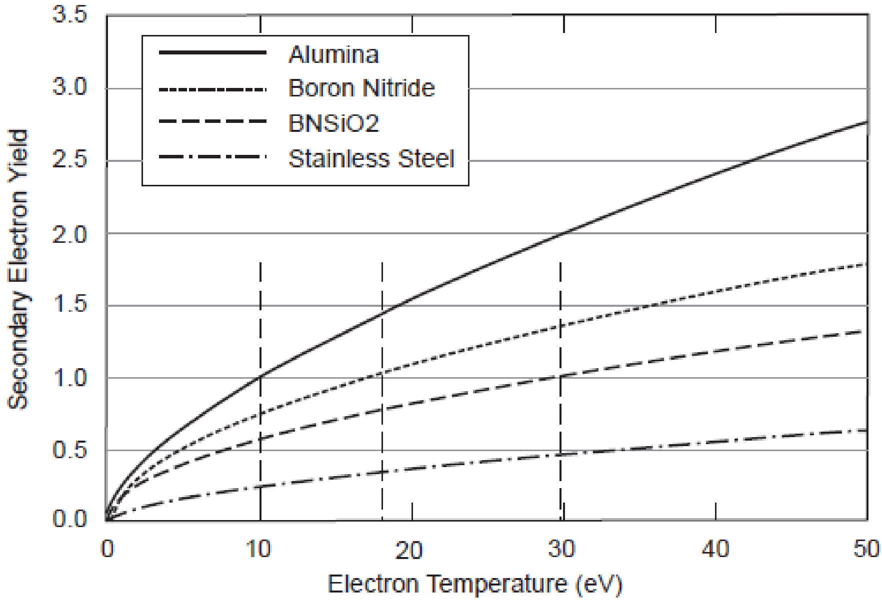


Fig. 1-10. Secondary electron yield for several wall materials used in Hall thrusters, measured with a mono-energetic electron beam

Measurements of the electron temperature in the channel of Hall thrusters by a number of authors [34–36] show electron temperatures in the channel well more than 20 eV. Equation (1.3-29) predicts that the sheath potential will go to zero and reverse from negative going (electron repelling) to positive going (electron attracting) as the secondary electron yield approaches unity for some of the materials. The value at which this occurs for each of the materials shown in Table 7-1 is indicated in Fig. 1-11. For boron nitride and alumina walls this occurs at electron temperatures below 20 eV, and for BN-SiO_2 walls it occurs at Electron temperatures on the order of 30 eV. In addition, depending on the collision mean-free path, some of the secondary electrons can pass completely through the plasma to strike the opposite wall of the channel. The possibility of the sheath potential reversing to electron attracting was used to predict very high electron power losses to the walls in some early analyses of Hall thrusters at high electron temperatures [30,31] because the incident electron flux can then equal or exceed the random electron flux along the magnetic field lines in the plasma.



Fig, 1-11. Secondary electron yield from the power-curve fits versus electron temperature, showing the cross-over value at which the yield equals one

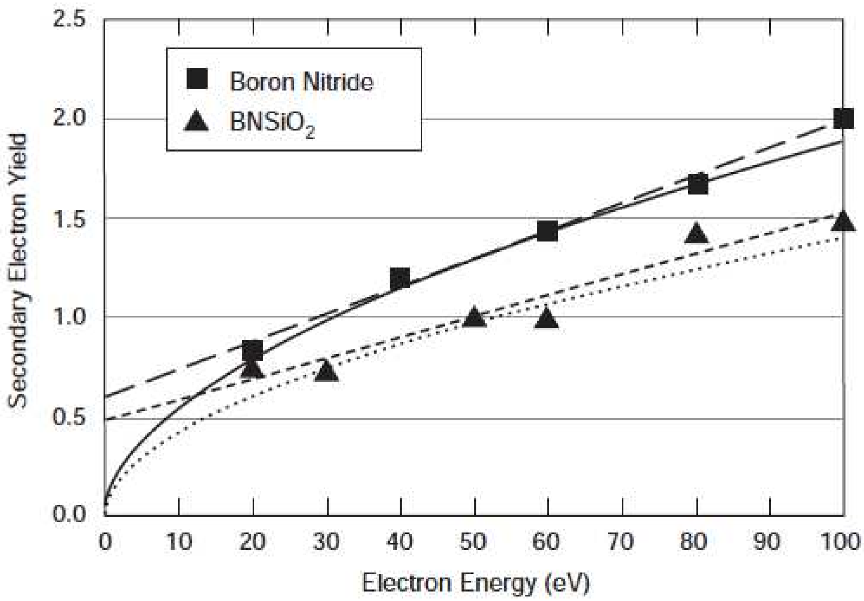


Fig. 1-12. Secondary electron yield versus electron energy, showing linear curve fits to the data producing finite yield at low incident energy

The sheath potential for a floating boundary can never go significantly more positive than the local plasma potential [37,38] for two reasons. First, the secondary electrons are ejected from the wall with very low energy (typically 1–2 eV). Any positive-going sheath (where the plasma is negative by one or two volts relative to the wall) will repel the secondary electrons and return them to the wall. This clamps the sheath potential to within a few volts positive with respect to the plasma. Second, the secondary electron emission is space charge–limited in the sheath. This effect was analyzed by Hobbs and Wesson [39], who showed that space charge limits the secondary electron current from the wall independently of the secondary electron yield. The local electron space charge in the sheath clamps the sheath voltage to a maximum value that is always negative relative to the plasma.

The effects of space charge on the sheath potential at the wall can be analyzed [39] By solving Poisson's equation for the potential in the sheath

$$\frac{\partial^2 \phi}{\partial x^2} = \frac{1}{\varepsilon_o}\left(n_e + n_s - n_i\right), \tag{1.3.31}$$

where n_s is the secondary electron density. Using a Maxwellian distribution for the electrons, the plasma density in the channel is

$$n_e = \left(n_o - n_{so}\right) e^{e\phi/kT}, \tag{1.3.32}$$

where n_o is the ion density at the sheath edge, n_{so} is the secondary electron density at the sheath edge, and is the potential relative to the potential $_o$ at the wall. The ions are assumed to be cold and to have fallen through the pre-sheath to arrive at the sheath edge with an energy of

$$\mathcal{E} = \frac{1}{2} m v_o^2, \tag{1.3.33}$$

where v_o is the Bohm velocity modified for the presence of the secondary electrons. The ion density through the sheath is then

$$n_i = n_o \left(\frac{\mathcal{E}}{\mathcal{E} - e\phi} \right)^{1/2}. \tag{1.3.34}$$

The secondary electrons are assumed to be emitted with an energy that is small compared to the plasma electron temperature and are accelerated through the sheath. The equation of continuity for current at the sheath edge gives,

$$n_s v_s = \frac{\gamma}{1-\gamma} n_o v_o, \tag{1.3.35}$$

where v_s is the secondary electron velocity. The secondary electron density through the sheath is then

$$n_s = n_o \frac{\gamma}{1-\gamma} \left(\frac{m}{M} \frac{\mathcal{E}}{\phi - \phi_o} \right). \tag{1.3.36}$$

Equations (7.3-32), (7.3-34), and (7.3-36) are inserted into Poisson's equation, Eq. (1.3-31), and evaluated by the usual method of multiplying through by d/dx and integrating to produce,

$$\begin{aligned} \frac{1}{2\varepsilon_o n_o kT_e} \left(\frac{d\phi}{dx} \right)^2 = {} & \frac{2\mathcal{E}}{kT_e} \left[\left(1 - \frac{e\phi}{\mathcal{E}} \right)^{1/2} - 1 \right] \\ & + \frac{2\gamma}{1-\gamma} \left(-\frac{m}{M} \frac{\mathcal{E}}{kT_e} \frac{e\phi_o}{kT_e} \right)^{1/2} \left[\left(1 - \frac{\phi}{\phi_o} \right)^{1/2} - 1 \right] \\ & + \left[1 - \frac{\gamma}{1-\gamma} \left(-\frac{m}{M} \frac{\mathcal{E}}{e\phi_o} \right)^{1/2} \right] \left[\exp\left(\frac{e\phi}{kT_e} \right) - 1 \right]. \end{aligned} \tag{1.3.37}$$

A monotonic sheath potential is found [39] for

For the case of no secondary electron emission (going to zero), the Bohm criteria solution of E = $kT_e /2e$ *is* recovered. Due to the large electron-to-ion mass ratio for xenon, the right-hand term is always small and the ion velocity at the sheath edge for the case of finite secondary electron emission will be near the Bohm velocity. Hobbs and Wesson evaluated this minimum ion energy at the sheath edge for the case of space charge–limited emission of electrons at the wall, d $_{o}$ $/ dx = 0$ in Eq. (7.3-37), and they found

$$E_o = 0.58 \frac{kT_e}{e}. \qquad (1.3.39)$$

Equation (7.3-39) indicates that the Bohm sheath criterion will still approximately apply (within about 16%) in the presence of secondary electron emission.

$$\frac{1}{4}\left[1-\frac{\gamma}{1-\gamma}\left(-\frac{m}{M}\frac{\mathcal{E}}{e\phi_o}\right)^{1/2}\right]\exp\left(\frac{e\phi_o}{kT_e}\right)\left(\frac{8kT_e}{\pi m}\right)^{1/2}=\frac{1}{1-\gamma}\left(\frac{2\mathcal{E}}{M}\right)^{1/2}$$

(1.3.40)

The value of the sheath potential for the space charge–limited case can be found by setting the electric field at the wall equal to zero in Eq. (7.3-37) and evaluating the potential using Eq. (7.3-38) and the current continuity equation:

The space charge–limited sheath potential for xenon is found to be

$$\mathcal{E}_o = 0.58\frac{kT_e}{e}.$$

(1.3.41)

The secondary electron yield at which the sheath becomes space-charge limited [39] is approximately

$$\gamma_o = 1 - 8.3\left(\frac{m}{M}\right)^{1/2},$$

(1.3.42)

which for xenon is 0.983.

This analysis shows that the sheath potential for a xenon plasma decreases from $5.97T_e$ for walls where the secondary electron yield can be neglected to $1.02T_e$ for the case of space charge–limited secondary electron emission that will occur at high plasma electron temperatures. The value of the sheath potential below the space-charge limit can be found exactly by evaluating the three equations, Eqs. (7.3-37), (7.3-38), and (7.3-40), for the three unknowns (and *E).*

However, the value of the sheath potential relative to the plasma edge in the presence of the secondary electron emission can be estimated by evaluating Eq. (7.3-29) while accounting for each of three species [38]. Quasi-neutrality for the three species in the plasma edge dictates that $n_i = n_e + n_s$, where n_s is the secondary electron density, and the flux of secondary electrons is the secondary electron yield times the flux of plasma electrons. Equating the ion flux to the net electron flux to the wall gives,

$$I_{iw} = n_i e v_i A = I_{ew}(1-\gamma) = \frac{1}{4} n_e (1-\gamma) e \left(\frac{8kT_e}{\pi m} \right)^{1/2} A \exp\left(\frac{e\phi_s}{kT_e} \right). \tag{1.3.43}$$

$$\phi_s = -\frac{kT_e}{e} \ln\left[\sqrt{\frac{M}{2\pi m}} \frac{n_e}{n_e + n_s} \frac{v_B}{v_i} (1-\gamma) \right]. \tag{1.3.44}$$

where the ion and electron densities are evaluated at the sheath edge. The sheath potential s relative to the plasma potential is then where vi is the modified ion velocity at the sheath edge due to the presence of the secondary electrons and the ion density is the sum of the plasma and secondary electrons. his equation is useful up to the space charge–limited potential of o = 1.02TeV and provides good agreement with the results for xenon described above for nevB / ni vi 0.5. The sheath potential predicted by Eq. (7.3-44) is plotted in Fig. 7-13 for two wall materials. In the limit of no secondary electron emission (= 0), the classic value for the sheath floating potential is obtained from Eq. (3.7-53). Once the electron temperature is sufficiently high to produce a yield approaching and even exceeding one, then the space charge–limited case of o = 1.02TeV is obtained. In between, the sheath potential depends on the electron temperature and material of the wall. Without the space charge–limited sheath regime predicted by Hobbs and Wesson, the potential would have continued along the thin dashed lines for the two cases and incorrectly resulted in very low sheath potentials and high-power loadings at the wall.

The total power to the wall of the Hall thruster is where the first term is due to electrons overcoming the repelling sheath potential and depositing $2T_e$ on the wall, and the second term is due to ions that have fallen through the pre-sheath potential and then the full sheath potential. Note that n_o in this equation is the plasma density at the sheath edge and is roughly half the average plasma density in the center of the channel due to the radial pre-sheath. The cooling of the wall by the secondary electron

$$P_w = \frac{1}{4}\left(\frac{8kT_e}{\pi m}\right)^{1/2} e n_o A e^{e\phi_s/kT_e}\left(2\frac{kT_e}{e}\right) + n_o e v_o A\left(\mathcal{E} - \phi_s\right).$$

(1.3.45)

Emission has been neglected. Equation (1.3-45) can be rewritten in terms of the total ion current to the wall as

$$P_w = I_{iw}\left[\left(\frac{M}{2\pi m}\right)^{1/2} e^{e\phi_s/kT_e}\left(2\frac{kT_e}{e}\right) + \left(\mathcal{E} - \phi_s\right)\right].$$

(1.3.46)

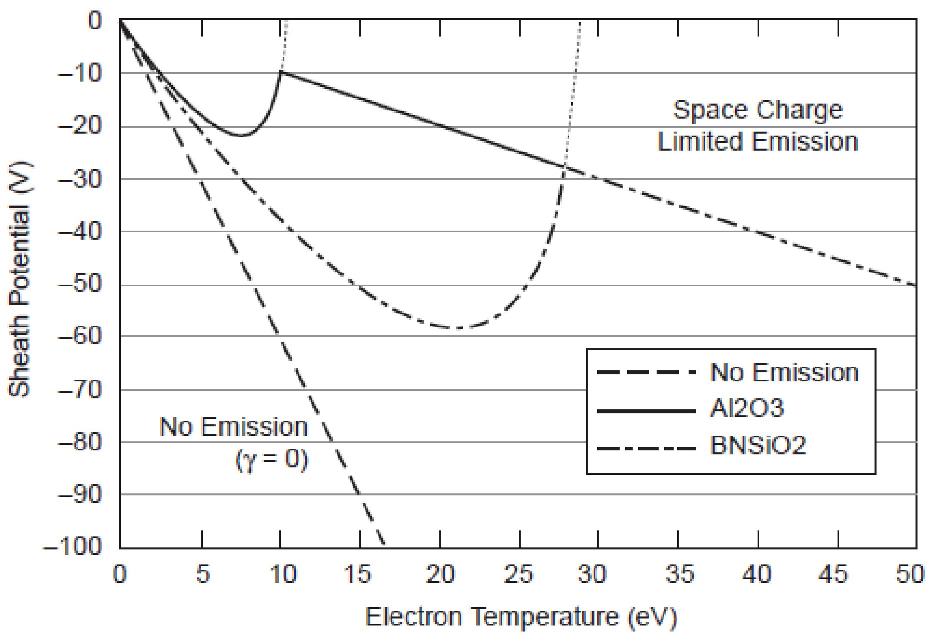


Fig. 7-13. Sheath potential versus electron temperature for two materials. The sheath transitions to space-charge limited where the dashed lines intersect the potential curves

For the case of space charge–limited secondary electron emission, the sheath potential is s = o = 1.02TeV , and the ion energy is E = 0.58 TeV in order to satisfy the Bohm condition. Equation (1.3-45) predicts the maximum heat loading to the wall in the presence of a Maxwellian electron distribution and secondary electron emission from the wall, which is the dominant power loss mechanism in dielectric-wall Hall thrusters. If the electron distribution function is non-Maxwellian, the heat load to the wall can differ from that predicted by Eq. (1.3-45).

In the case of TAL thrusters, the channel wall is metallic and biased to the cathode potential. This eliminates the zero-net current condition found on the insulating walls of dielectric-channel Hall thrusters and used to determine the local heat flux in Eq. (1.3-45). The electron flux to the cathode-biased TAL channel wall is negligible, and the secondary yield for metals is much lower than for insulators, so the secondary electron emission by the wall in TAL thrusters has little effect on the thruster operation. In addition, the plasma tends to be localized near the channel center by the anode design and gas feed geometry. The plasma then tends to be in poor contact with the guard rings at the wall that also have a small exposed area to the plasma, resulting in low radial ion currents to the wall. This is evidenced by the erosion pattern typically observed on TAL guard rings [29], which tends to be on the downstream face from particles outside the thruster instead of on the inside diameter from the channel plasma. While the ion and electron currents and power deposition to the inside diameter of the metallic guard ring are likely smaller than in the dielectric-wall thruster case (where the power loss due to the electrons is dominant), the erosion on the face of the guard ring indicates energetic ion bombardment is occurring. This effect is significant in determining the life of the TAL.

However, TAL thrusters are characterized by having the anode in close contact with the magnetized plasma near the channel exit, in contrast to the dielectric- wall Hall thrusters. The magnetized plasma has a high electron temperature, which causes a significant amount of power to be deposited from the discharge current on the anode. It is possible to evaluate this power loss mechanism based on the current and sheath potential at the anode. As described above, the discharge current is essentially equal to the electron current collected at the anode. In order for the TAL thruster to transfer a large fraction of the discharge voltage to the ions, the potential of the plasma near the anode must be close to the anode potential. Assuming the local plasma potential is then equal to or slightly positive relative to the anode, the electron current to the anode, I_a , deposits $2T_{\mathrm{eV}}$ in energy from the plasma (see Appendix C). The power deposited on the anode, P_a , is then given by

$$P_a = 2T_{\mathrm{eV}} I_a \approx 2T_{\mathrm{eV}} I_d , \qquad (1.3.47)$$

where Eq. (7.2-26) has been used. If the plasma potential is negative relative to the anode, the thruster efficiency will suffer due to the loss of discharge voltage available to the ions, and the anode heating will increase due to the positive- going sheath potential accelerating electrons into the anode. Equation (1.3-47) then represents a reasonable, but not worst-case, heat flux to the anode.

This power loss to the anode can be related to the beam current using the fraction of the discharge current that produces beam current, which is defined as

$$\eta_b = \frac{I_b}{I_d} . \qquad (1.3.48)$$

Therefore, the power to the anode is in well-designed Hall thrusters, b ranges typically from 0.6 to 0.8. Therefore, the power loss to the anode is 3 to 4 times the product of the electron temperature in the near-anode region and the beam current. This is the most significant power loss mechanism in TAL thrusters.

$$P_a = 2T_{\mathrm{eV}} \frac{I_b}{\eta_b} . \qquad (1.3.49)$$

7.3.4 Plasma Electron Temperature

The electron temperature in the channel must be known to evaluate the power loss mechanisms described above. The peak electron temperature in the plasma channel can be found using power balance, described by Eq. (1.3-27). This method provides reasonable estimates because the power loss in the thruster will be shown to be a strong function of the electron temperature. Even though the plasma density and electron temperature peak in different locations along the channel associated with the different ionization and acceleration regions, the strong axial electron temperature profile in Hall thrusters causes the majority of the power loss to occur in the region of the highest electron temperature. This occurs near the channel exit where the magnetic field across the channel is the strongest. Evaluating the plasma parameters and loss terms in this region, which is bounded by the channel width and magnetic axial field extent in the channel, establishes the electron temperature that is required to satisfy the power balance in the plasma for a given thruster current and voltage.

The individual terms in Eq. (1.3-27) will now be evaluated. The input power to the thruster is the discharge current times the discharge voltage ($P_d = I_d V_d$). The power in the beam, using Eq. (1.3-48), is

$$P_b = \eta_b \eta_v I_d V_d = \eta_v I_b V_d, \tag{1.3.50)}$$

where the current utilization and voltage utilization efficiencies have to be known or evaluated by some means. The difference between the beam power and the discharge power is the power remaining in the plasma channel to produce the plasma and offset the losses:

$$P_p = (1-\eta_b) I_d V_d = I_{ec} V_d, \tag{1.3.51}$$

Where P_p is the power into the plasma. The plasma is produced and heated essentially by the collisional transport of the electrons flowing from the cathode plasma in the near-plume region to the anode inside the thruster. The power into channel walls, from Eq. (7.3-45), can be written as

$$P_w = n_e e A \left[\frac{kT_e}{e} \left(\frac{kT_e}{2\pi m} \right)^{1/2} e^{e\phi_s / kT_e} + \frac{v_i}{2} (\varepsilon - \phi_s) \right], \tag{1.3.52}$$

Where A is the total area of the inner and outer channel walls in contact with the high temperature plasma region, v_i is the ion velocity toward the wall, and the sheath potential $_s$ is given by Eq. (7.3-44). Equation (7.3-52) shows the wall power varies linearly with density but with the electron temperature to the 3/2 power. Therefore, the dominant wall losses occur in the region of the highest electron temperature.

The power into the anode, from Eq. (7.3-47), can be written as

$$P_a = 2I_d T_{eV}(\text{anode})\,. \tag{1.3.53}$$

Where the electron temperature in this case is evaluated near the anode. The power radiated is

$$P_R = n_o n_e \langle \sigma_* v_e \rangle V, \tag{1.3.54}$$

where the excitation reaction rate coefficient is given in Appendix E as a function of the electron temperature, and *V* is the volume of the high- temperature plasma region in the channel, which can be taken to be the channel cross-sectional area times the axial thickness *L*. Equations (1.3-52) and (1.3-54) require knowledge of the plasma density in the high-temperature region in the channel. This can be found to first order from the beam current

$$n_e = \frac{I_b}{e v_b A_c} \approx \frac{\eta_b I_d}{e A_c \sqrt{\dfrac{2\eta_b e V_d}{M}}}, \tag{1.3.55}$$

where A_C is the area of the channel exit. Finally, the power to produce the ions in the thruster is the sum of the beam current and the ion current to the walls times the ionization potential:

$$P_{\text{ion}} = \left(I_b + I_{iw}\right)U^+ = \left[\eta_b + I_{ew}\left(1-\gamma\right)\right] I_d U^+\,. \tag{1.3.56}$$

The peak electron temperature is found by equating the input power to the plasma in Eq. (7.3-51) with the sum of the various loss terms described above, and then iterating to find a solution. For example, the SPT-100 Hall thruster has a channel outside diameter of 10 cm, a channel inside diameter of 7 cm, and runs nominally at a discharge of 300 V at 4.5 A with a current utilization efficiency of 0.7 and a voltage utilization efficiency of 0.95 [6]. From Eq. (7.3-55), the plasma density at the thruster exit is about 1.6 10^{17}m^{-3}. The power into the plasma, from Eq. (7.3-51), is about 433 W. Taking the electron temperature at the anode to be 5 eV and the hot-plasma thickness *L* to be about

1 cm, the power balance equation is satisfied if the electron temperature in the channel plasma is about 25 eV.

It is a common rule-of-thumb in Hall thrusters to find that the electron temperature is about one-tenth the beam voltage [35]. The result in the example above of T_e 0.08 V_d is consistent with that observation. It is also important to note that nearly 70% of the power deposited into the plasma goes to the dielectric channel walls in the form of electron heating, and that the radiation losses predicted by Eq. (7.3-54) are negligible for this case because the electron temperature is so high. Finally, the ion current to the wall for this example from the solution to Eq. (7.3-28) is 0.52 A, which is about 12% of the discharge current and 8% of the beam current in this thruster. This amount agrees well with the 10% of the ion current going to the wall calculated by Baranov [40] in analyzing Hall thruster channel wear.

where I_{iw} is given by Eq. (7.3-28) and I_{ew} is given by the left-hand side of Eq. (7.3-52) divided by $2T_e$ (because the electron energy hitting the wall is already included in this equation).

7.3.5 Hall Thruster Efficiency (Dielectric Walls)

The efficiency of a Hall thruster with a dielectric wall can be estimated by evaluating the terms in the thruster efficiency given which requires evaluating the total power-loss terms in Eq. (7.3-27) to obtain a value for the effective electrical efficiency. This also illustrates the dominant loss mechanisms in the thruster.

The first term in Eq. (1.3-27), the beam power due to the accelerated ions, P_b , is just $I_b V_b$, where the effective beam voltage will be used. The power loss to the dielectric wall will be estimated for the SPT-100 Hall thruster [4–6] using the analysis of Hobbs and Wesson [39] described in Section 1.3.3. The heat flux to the wall was given by Eq. (1.3-46):

$$P_w = I_{iw}\left[\left(\frac{2M}{\pi m}\right)^{1/2} e^{e\phi_s/kT_e}\left(\frac{kT_e}{e}\right)+\left(\mathcal{E}-\phi_s\right)\right],$$

(1.3.57)

where I_{iw} is the ion flux to the wall. Following Hobbs and Wesson, the modification to the Bohm criterion is small and $E\ T_e/2$ from the Bohm criterion. From Eq. (7.3-44), the sheath potential for xenon and BNSiO2 walls in the SPT-100 thruster, assuming an average electron temperature along the channel wall of 25 eV, is about –54 V. Plugging these values into Eq. (7.3-57) gives

$$P_w = 45.8 I_{iw} T_{eV} + 2.65 I_{iw} T_{eV} = 48.5 I_{iw} T_{eV} \quad (1.3\text{-}58)$$

The first term on the right-hand side is again the electron power loss to the wall (written in terms of the ion current to the dielectric surface), and the second term is the ion power loss. The power loss to the channel wall due to the electron loss term is an order of magnitude larger than the power loss due to ions.

It is convenient in evaluating the efficiency of the thruster to relate the ion current to the wall in Eq. (1.3-58) to the beam current. In the plasma, there is an electric field toward the wall due to the pre-sheath of approximately $T_{eV}/2r = T_e/w$. There is also the axial electric field of V_b/L producing the beam energy. It is common in Hall thrusters to find that the electron temperature is about one-tenth the beam voltage [35], and the channel width is usually approximately L [4,20]. Therefore, the axial electric field is on the order of 10 times the radial electric field. On average, then, the ion current to the channel walls will be about 10% of the beam current. This very simple argument agrees with the SPT-100 example results given in the previous section and the results of Baranov [40]. Using Eq. (1.3-58) with the above estimates for the ion current and electron temperature, the power loss to the insulator walls is

$$P_w = 48.5 I_{iw} T_{eV} = 48.5(0.1 I_b)(0.1 V_b) = 0.49 I_b V_b \,. \quad (1.3\text{-}59)$$

The power loss to the anode is due to the plasma electrons overcoming the

sheath potential at the anode surface. From Eq. (7.2-24), the anode electron current is

$$I_{ea} = I_d + I_{ia} \,. \qquad (1.3\text{-}60)$$

Neglecting the ion current to the anode as small (due to the mass ratio), and realizing that each electron deposits $2kT_e/e$ to the anode for positive plasma potentials (from Appendix C), the power to the anode is

$$P_a = 2T_{\mathrm{eV}} I_d . \qquad (1.3\text{-}61)$$

The electron temperature near the anode is very low, typically less than 5 eV [34–36]. Using the thruster current utilization efficiency and assuming

$_b = 0.7$ and $T_{\mathrm{eV}} = 0.01V_b$ near the anode, this can be written as

$$P_a = 2 {}_b I_b (0.01V_b) = 0.014 I_b V_b . \qquad (1.3\text{-}62)$$

The power required to produce the ions is given by Eq. (7.3-56). This can be written as

$$P_{\mathrm{ion}} = (I_b + I_{iw})U^+ = (1 + {}_b) I_d U^+ \qquad (1.3\text{-}63)$$

Taking the beam utilization efficiency as 0.7 and estimating that the ionization potential is roughly 5% of the beam voltage, the power required to produce the ions is approximately $P_i = 0.09\ I_d V_b$. The radiation power and other power loss mechanisms are small and will be neglected in this simple example.

The total discharge power into the thruster is then

$$P_d = I_b V_b + 0.49\ I_b V_b + 0.014\ I_b V_b + 0.09\ I_b V_b = 1.59\ I_b V_b . \quad (1.3\text{-}64)$$

The electrical efficiency of the dielectric-wall thruster is then

$$_e = I_b V_b / (1.59\ I_b V_b) = 0.63 . \qquad (1.3\text{-}65)$$

The total thruster efficiency, assuming the same beam divergence and double- ion content as evaluated above and a mass utilization efficiency of 95% reported for SPT thrusters [4], is

$$_T = (0.915)^2\ (0.63)(0.95) = 50\% . \qquad (1.3\text{-}66)$$

The SPT-100 thruster is reported to run at about 50% efficiency. Since the power loss is dominated by the electron wall losses, this analysis illustrates how critical the wall material selection is to minimizing the secondary electron yield and maintaining a sufficient wall sheath potential for good efficiency. For example, if the wall had been made of alumina and the electron temperature was about 20 V, the sheath potential would be $-1.02T_{\mathrm{eV}}$ in the space charge–limited regime. The wall power from Eq. (1.3-57) would then be about three times higher than in the BNSiO$_2$ case:

$$P_w = 142 I_{iw} T_{\mathrm{eV}} = 1.4 I_b V_b . \qquad (1.3.67)$$

The electrical efficiency of the thruster, assuming the same anode loading and energy loss to the beam, would be, $_e\ 0.40$ and the total efficiency would be

$$_T = (0.915)^2\ (0.40)(0.95)\ 32\% \quad (1.3\text{-}68)$$

Recent parametric experiments in which different wall materials were used in the SPT-100 [33] showed that changing from $BNSiO_2$ to alumina reduced the efficiency to the order of 30%, consistent with the increased secondary electron yield of the different wall material.

The agreement of this simple analysis with the experimentally measured efficiencies is somewhat fortuitous because the predictions are very sensitive to the secondary electron yield of the wall material and the actual sheath potential. Small errors in the yield data, changes in the wall material properties during thruster operation, and inaccuracies in the empirical values for the electron temperature and ion flux with respect to the beam parameters will significantly affect the calculated results. Other effects may also be significant in determining the thruster efficiency. The analysis of the sheath potential assumed a Maxwellian electron distribution function. It was recognized several years ago [37,41,42] that the electron distribution may not be Maxwellian. Detailed kinetic modeling of the Hall thruster channel plasma [43,44] indicates that the electron velocity distribution is depleted of the high-energy tail electrons that rapidly leave the plasma along the magnetic field lines and impact the wall. This is especially true near the space-charge limit where the sheath voltage is small and a large fraction of the electron tail can be lost. The collision frequencies and thermalization rates in the plasma may be insufficient to re-populate the Maxwellian tail. This will effectively result in a lower electron temperature in the direction parallel to the magnetic field toward the walls [45], which can increase the magnitude of the sheath potential and reduce the electron heat loss to the wall. In addition, re-collection of the secondary electrons at the opposite wall [46,47], due to incomplete thermalization of the emitted secondary electrons in the plasma, modifies the space-charge limits and sheath potential, which also can change the electron heat flux to the wall.

These effects are difficult to model accurately due to the presence of several different electron populations, several collision/thermalization processes, the effect of magnetization on the electrons, and the presence of plasma instabilities. Understanding what determines the electron temperature and velocity distribution as a function of the discharge voltage and current, and uncovering the effects that determine the wall power flux and finding techniques to minimize them, are continuing areas of research now.

7.3.6 TAL Hall Thruster Efficiency (Metallic Walls)

As with the 1.35-kW SPT-100 Hall thruster example above, an estimate will be made of the power loss terms in Eq. (7.3-27) to obtain an electrical efficiency for the 1.4-kW D-55 TAL thruster [29]. will then be used to obtain an estimate for the thruster efficiency. The beam power P_b is, again, just $I_b V_b$. As stated in the previous section, the wall losses (P_w) are essentially negligible in TAL thrusters, and the power to the anode is given by Eq. (1.3-49):

$$P_a = 2T_{\text{eV}} \frac{I_b}{\eta_b} = 0.29 I_b V_b \, . \qquad (1.3.69)$$

In Eq. (7.3-69), it is again assumed $_b$ = 0.7 and $T_{\text{eV}} = 0.1V_b$, although these values may be somewhat different in TAL thrusters. The power to produce the ions is again approximately $0.09 I_b V_b$.

The total discharge power, Eq. (7.3-27), then becomes

$$P_d = I_b V_b + 0.29\ I_b V_b + 0.09\ I_b V_b = 1.4\ I_b V_b \, . \quad (1.3\text{-}70)$$

Neglecting the power in the cathode keeper (if any) and the magnet as small compared to the beam power, the electrical utilization efficiency from Eq. (2.5-1) is then

$$\eta_e = \frac{P_d}{1.4 P_d} = 0.72 \, .$$

(1.3.71)

The total thruster efficiency, assuming a 10% double-ion content, a 20-deg angular divergence [29,48], and a 90% mass utilization efficiency reported for TAL thrusters [29,49], is then, from Eq. (2.4-7),

$$_T = (0.915)^2\ (0.72)(0.9) = 54\% \, . \quad (1.3\text{-}72)$$

This result is on the same order as that reported in the literature [29,49,50] for this power-level TAL and is essentially the same as the SPT-100 efficiency in this simple example if the wall losses had been included. However, the power loss to the anode is the dominant mechanism in the TAL efficiency.

7.3.7 Dielectric-Wall Versus Metallic-Wall Comparison

It is interesting to make a few direct comparisons of dielectric-wall Hall thrusters with metallic-wall TAL thrusters. Similar discussions have appeared in the literature [1,3,31], often with conflicting opinions. The basic plasma physics in the channel described above applies to both the dielectric-wall Hall thruster and the TAL.

The maximum electron temperature occurs in both thrusters near the channel exit in the region of strongest magnetic field where the Hall current is a maximum. The different interaction of the thruster walls with this plasma determines many of the characteristics of the thruster, including life. Dielectric-wall thrusters have a significant amount of their input power deposited as loss on the dielectric channel walls due to electron bombardment. In the above example efficiency calculation, approximately 25% of the power going into the thruster was deposited on the channel walls. The metallic walls in TAL thrusters collect a smaller electron current because they are biased to cathode potential, and they also tend to have a small exposed area in poor contact with the plasma, which limits the amount of ion and power lost to these surfaces. However, the anode is positioned very close to the high electron temperature region and receives a significant amount of power deposition in collecting the discharge current. In the above example TAL efficiency calculation, over 20% of the power going into the thruster was deposited on the anode.

The deep channel in dielectric-wall Hall thrusters, with a low magnetic field strength and low electron temperature near the anode, tends to minimize the power deposition on the anode. In the above simple example, only 1% of the thruster input power was deposited on the anode. Nevertheless, the anode is normally electrically isolated from the thruster body (and therefore thermally isolated), and so anode overheating is sometimes an issue, especially at high power density. The anode in TAL thrusters can also have heating issues because the loading is much higher, even though the view factor for the anode to radiate its power out of the thruster is better than the deep channel in the insulting-wall configuration. In addition, with the anode positioned physically close to the thruster exit in TALs, impurity deposition and material buildup problems can occur. This has been an issue in ground testing of some TAL thrusters [29], where carbon deposition on the anode from back sputtering from the beam dump became significant over time. TAL thrusters with deeper channels can be designed and operated [3]. The performance of the thruster is likely different in this configuration, and ion bombardment and sputtering of the metallic channel walls can become significant and affect the thruster life.

Dielectric-wall Hall thrusters are often described in terms of an ionization zone upstream of the exit plane and an acceleration zone in the region of the exit plane. TAL thrusters have a similar ionization region near the magnetic field maximum, which is now closer to the anode because the magnetic field gradient is greater. The TAL acceleration zone is a layer close to the anode [1,3] that can extend outside of the thruster [48]. The higher electron temperatures associated with TAL thrusters support higher electric fields in the quasi-neutral plasma, which compresses these zones relative to dielectric-wall thrusters. In addition, the metallic walls and higher electric fields are conducive to multiple acceleration stages, which can improve thruster performance and produce higher Isp than a conventional single-stage TAL thruster [1,51]. Multiple-stage dielectric-wall Hall thrusters that operate at high Isp have also been investigated (see [17] and the references cited therein).

Finally, the difference between dielectric-wall Hall thrusters and TAL thrusters is sometimes attributed to the secondary electron coefficients of the different wall materials. The above discussion shows that this is not the dominant difference. Instead, the proximity of the TAL anode electrode to the high temperature plasma region and the thruster exit plane is what changes the electric field profile, power deposition, and sputtering characteristics as compared to the dielectric-wall Hall thruster.

7.4 Channel Physics and Numerical Modeling

As discussed in the previous sections, the detailed physics determining Hall thruster performance is not well understood. Specifically, the electron distribution function in the exit region, the mechanisms responsible for electron transport across the magnetic field, and the role of oscillations on the particle transport and plasma conditions need to be determined. A considerable effort has been made to develop fluid, kinetic, hybrid, and particle-in-cell (PIC) models to predict and explain the performance and effects observed in Hall thrusters.

Hirakawa and Arakawa developed [52] a two-dimensional (2-D) particle-in-cell model where anomalous electron diffusion was introduced by using oscillating azimuthal electric fields. Boeuf and Garrigues developed a one-dimensional (1-D) hybrid model [53] in which the electrons were treated as a fluid and the ions were described by a collision less Vlasov (kinetic) equation. Similar fluid and hybrid models have been developed by other authors [54–56] using various techniques to determine the ion transport, such as Monte-Carlo simulation, Boltzman equation solutions, and “ion free-fall” (essentially a Bohm current solution) to the boundaries. The most widely used code, HPHall, is a 2-D, transient hybrid model originated by Fife and Martinez-Sanchez [30] that has been recently extended with an improved sheath model [41,42,57] and a model of channel erosion.

7.4.1 Hybrid Hall Thruster Models

Hybrid Hall thruster models, such as HPHall [30, 57], utilize a steady-state fluid electron momentum equation and a time-dependent electron energy equation to solve for electron temperature and potentials in the channel and plume. The codes also use time-dependent ion and neutral particle equations to calculate the plasma density and ion velocities on a time scale much larger than the electron time scale. These codes are also used to model Hall thruster transit- time oscillations that are on the order of time scales related to neutral atom and ion motions (1 MHz) but cannot capture the effects of electron instabilities that have much higher frequencies.

From the steady-state electron momentum equation, an Ohm’s law representation from Eq. (3.6-20) for the electron field is

$$\eta \mathbf{J}_e = \mathbf{E} + \frac{\nabla \mathbf{p} - \mathbf{J}_e \times \mathbf{B}}{en} - \eta_{ei} \mathbf{J}_i \, . \qquad (1.4.1)$$

where the resistive term has the following form in the magnetic frame of reference:

$$\eta \mathbf{J} = \eta_{\perp} J_{\perp} + \eta_{\parallel} J_{\parallel} + \eta_{\wedge} J_{\wedge} \, , \qquad (1.4.2)$$

and the subscripts represent the directions perpendicular, parallel, and transverse (in the **E B** direction), respectively, to the local magnetic field.

Equation 7.4-1 must be separated into the two components of the $\mathbf{J}_e \mathbf{B}$ motion in a manner like that in Section 3.6 and solved for the electric field. From current conservation, the electron current is taken to be the difference between the discharge current and the ion current from the particle calculations. Typically, the circuit current is chosen at each time step to satisfy the applied voltage (= **E d**) boundary conditions.

7.4.1.1 Transverse Electron Transport.

Writing the perpendicular resistivity in terms of the perpendicular electron mobility, as defined in Eq. (3.6-66), gives

$$\eta_{\perp} = \frac{1}{en\mu_{e\perp}} = \frac{1+\omega_c^2\tau_m^2}{en\mu_e} = \frac{1+\omega_c^2/\nu_m^2}{en\mu_e},$$

(1.4.3)

where the collision time m for momentum transfer is equal to one over the collision frequency (1/ m). The perpendicular electron flux from Ohm's law, Eq. (7.4-1), can then be written as

$$J_{e\perp} = \mu_{e\perp}\left(enE_{\perp} + \frac{\partial p_e}{\partial x}\right) - \frac{\mu_{e\perp}}{\mu_{ei}} J_{i\perp},$$

1.4.4)

and the electron mobility due only to electron–ion collisions is given by m

Usually, the ion flux term in Eq. (1.4-4) is neglected and an effective electric field is used such that the electron flux is expressed as

$$\mu_{ei} = \frac{e}{m\nu_{ei}}.$$

(1.4.5)

$$J_e = en\mu_e E, \qquad (1.4\text{-}6)$$

expression for the transverse electron mobility then accounts for both electron–ion and electron–neutral collisions in the partially ionized plasma. Since the electrons are well magnetized in the plasma near the exit of the channel where the magnetic field strength is the highest, the electron Hall parameter is much greater than unity and the transverse electron mobility across the field lines is found to be small. In fact, calculations of the electron collision frequency based on the classical collision terms in Eq. (7.4-9) are unable to provide sufficient cross-field transport to support the discharge current passing through the thruster [54,57,59]. In addition, the neutral density in the plume of the Hall thruster is low due to the high mass utilization efficiency, which reduces the effective collision frequency in Eq. (7.4-9) and again leads to problems in providing sufficient transport of the electrons from the external.

Where the effective electric field is

$$E'_{\perp} = \left(E_{\perp} + \frac{1}{en} \frac{\partial p_e}{\partial x} \right). \tag{1.4.7}$$

The effective perpendicular electron mobility in Eq. (1.4-4) is

$$\mu_{e\perp} = \frac{e}{m\nu_m} \frac{1}{1 + \omega_c^2 / \nu_m^2} = \frac{\mu_e}{1 + \Omega_e^2}, \tag{1.4.8}$$

where Ω^2 - is the electron Hall parameter, and the momentum-transferring collision frequency ν_m is described

$$\nu_m = \nu_{ei} + \nu_{en}. \tag{1.4.9}$$

This expression for the transverse electron mobility then accounts for both electron–ion and electron–neutral collisions in the partially ionized plasma.

Since the electrons are well magnetized in the plasma near the exit of the channel where the magnetic field strength is the highest, the electron Hall parameter is much greater than unity and the transverse electron mobility across the field lines is found to be small. In fact, calculations of the electron collision frequency based on the classical collision terms in Eq. (7.4-9) are unable to provide sufficient cross-field transport to support the discharge current passing through the thruster [54, 57, and 59]. In addition, the neutral density in the plume of the Hall thruster is low due to the high mass utilization efficiency, which reduces the effective collision frequency in Eq. (1.4-9) and again leads to problems in providing sufficient transport of the electrons from the external cathode across the transverse field lines and into the channel to support the discharge current. Two mechanisms have been proposed in an attempt to describe "enhanced" cross-field electron transport and explain the observed Hall thruster operation. Morozov [12] postulated that electron-wall interactions in the channel region will scatter electron momentum and introduce secondary electrons, which can increase the effective cross-field transport. This effect is introduced into the effective collision frequency by a wall-scattering frequency $_{w}$:

$$\nu_m = \nu_{ei} + \nu_{en} + \nu_w \,. \tag{1.4.10}$$

The wall-scattering frequency is either given by 10^7 per second [53], with an adjustable parameter used to match the experimental data, or the wall collision frequency of electrons is calculated directly in the code [59]. While this effect does increase the electron transport in the channel, it is sometimes found to provide insufficient enhancement of the electron transport. In addition, in the plume of the thruster there are no walls and the neutral density is very low, which precludes the use of Eq. (1.4-10) to increase the cross-field transport sufficiently to explain the experimental data.

Additional cross-field transport has been added in the codes by invoking Bohm diffusion both inside and outside the thruster channel. As discussed in Chapter 3, Bohm diffusion likely arises from **E B** driven drift instabilities, which can naturally occur in these thrusters due to the Hall current. Using the Bohm diffusion coefficient from Eq. (3.6-72) and the Einstein relationship of Eq. (3.6-28), a Bohm mobility can be defined as

$$\mu_B = \frac{1}{\beta B} = \frac{e}{\beta m \omega_c} \,, \tag{1.4.11}$$

where is an adjustable coefficient changed to make the code predictions of the thruster parameters fit the experimental data. If full Bohm diffusion is required by the code to match the data, such as is often the case in the plume, then = 16. The effective Bohm collision frequency is then

$$\nu_B = \beta \omega_c \,. \tag{1.4.12}$$

The total "anomalous" collision frequency used in the codes is

$$\nu_m = \nu_{ei} + \nu_{en} + \nu_w + \nu_B \,, \tag{1.4.13}$$

Where the wall collision frequency $_w$ is neglected in the plume.

7.4.1.2 Transport along the Magnetic Field.

In the direction along the magnetic field lines, the J B cross product in the electron momentum equation is zero and Eq. (1.4-1) becomes

$$\eta J_e = E + \frac{\nabla p}{en} - \eta_{ei} J_i \,. \tag{1.4.14}$$

The electric field along the field line is then

$$E = \frac{\nabla p}{en} + \eta J_e - \eta_{ei} J_i = -\nabla \phi \,. \tag{1.4.15}$$

With the standard assumptions used along magnetic fields in many plasmas of zero net current ($j_e\ j_i$) and uniform electron temperature, Eq. (7.4-15) can be solved for the potential along the field line to give

$$\phi = \phi_o + T_e \ln\left(\frac{n}{n_o}\right). \tag{1.4.16}$$

This equation was derived in Section 3.5-1 and represents the simple Boltzman relationship for plasmas with Maxwellian electron distribution functions. It is often called the *barometric law* in ion thruster literature and the *thermalized potential* in Hall thruster literature. Thus, the transport along the magnetic field lines is usually considered to be classical.

It is commonly assumed that the density gradient along the magnetic field line is relatively small, so the potential change along a magnetic field line from

Eq. (7.4-16) is essentially zero. Therefore, within about kT_e /e , the magnetic field lines represent equipotential lines in the plasma. The simplifying assumptions leading to this conclusion (zero net current, Maxwellian electrons, and small density gradient along the magnetic field lines) are often used and may introduce significant errors in some cases. Nevertheless, the thermalized potential has been used for many years [3] in the design of Hall thrusters to relate the magnetic field shape to the electric field in the plasma [11].

7.4.1.3 Ion Current.

Several methods have been used to describe the ion generation and transport in the Hall thruster models. First, the ions have been modeled as a fluid using continuity equations [54,60], where the axial motion is due to the electric field along the channel and the radial motion to the wall is determined by the ion-neutral scattering frequency. The ion current to the wall is then

$$I_W = n_i n_o {}_{\text{in}} v_i A_W L , \qquad (1.4\text{-}26)$$

where A_W is the wall area, L is the plasma length, $_{\text{in}}$ is the ion-neutral collision cross section for 90-deg scattering including elastic and charge- exchange collisions, and the velocity of the neutrals is neglected relative to the ion velocity. In PIC numerical codes, this represents the radial flux to the cell boundary where $A_W L$ becomes the cell volume.

Fife [30] modeled the ion motion using a 2-D PIC code that assumed the ions and neutrals acted as discrete macro-particles in each cell. The time step in the ion-PIC code, in this case, was adjusted (to typically three orders of magnitude slower than the electron model time step) to handle the ion-motion time scales without invoking excessive computational time.

Finally, the ion Vlasov equation has been used to solve for the ion generation and motion [52,53]. This has primarily been applied for investigating low- frequency oscillations on the order of the ion-characteristic time scales. In one dimension, this can be written as

$$\frac{\partial f}{\partial t} + v_x \frac{\partial f}{\partial t} + \frac{e}{M} E \frac{\partial f}{\partial v_x} = n_e n_o \langle \sigma_i v_e \rangle \delta(v_x - v_o). \qquad (1.4.27)$$

where f is the ion distribution function and $(v_x \; v_o)$ is the Dirac delta function valuated for the ion velocity relative to the neutral velocity. The ion density is then found from

$$n_i = f(x, v_x, t) dv_x . \qquad (1.4\text{-}28)$$

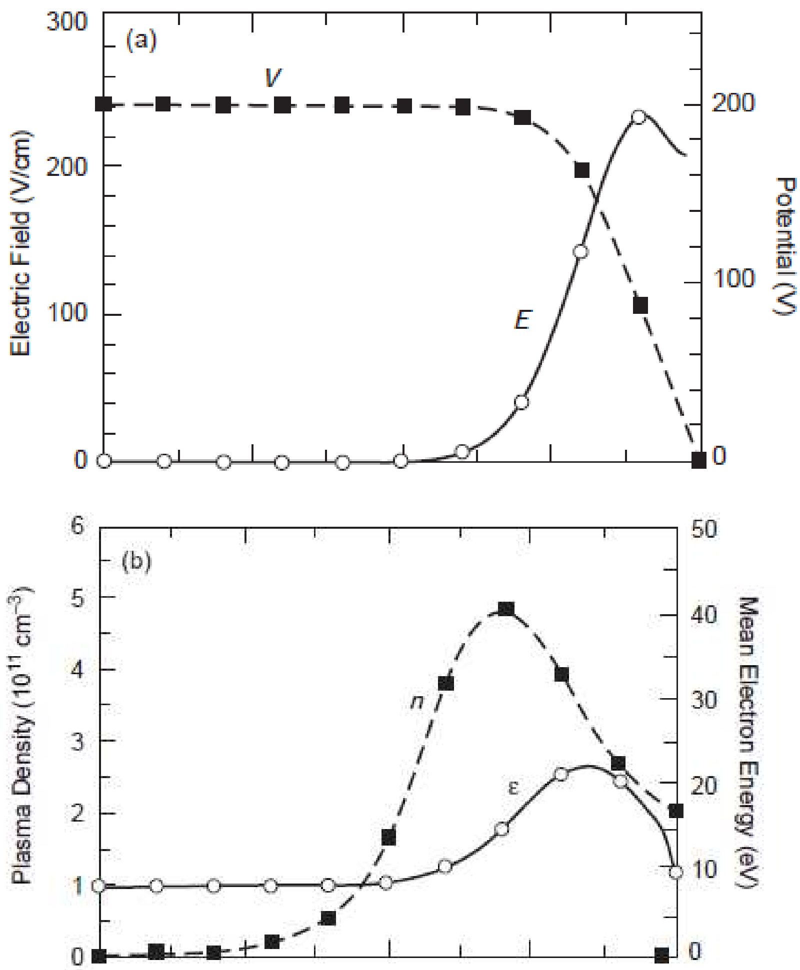
The plasma is always assumed to be quasi-neutral ($n_i \; n_e$). At the sheath boundary at the wall, the ion current normally is assumed to be the Bohm current and the electron current is the one-sided random electron flux. Total current continuity requires the ion flux and net electron flux (incident electrons and emitted secondary electrons) to the insulating walls to be equal, which establishes the sheath potential to produce quasi neutrality and charge conservation as described above. The hybrid-model equations described above for determining the ion currents are normally evaluated numerically in either 1-D or 2-D with greatly different time steps between the electron fluid evolution and the ion and neutral motion evaluations.

7.4.2 Steady-State Modeling Results

The physics of the Hall thruster discharge related to the transverse electron mobility, electron-wall interactions, and the exact nature of the electron distribution function are not completely understood at this time. However, the 1-D and 2-D models described above are reasonably successful in predicting plasma parameters and thruster behavior provided enhanced electron conductivity is incorporated in the channel due to wall collisions and turbulence, and modifications to the wall heat fluxes are made associated with the secondary electron behavior. In addition, enhanced electron transport in the plume region near the thruster exit is required to match the models' predictions with the experimental results [61], which is normally provided by assuming collective oscillations drive Bohm-like diffusion. In this region, other mechanisms may also be responsible for the cross-field electron transport, and research in this area to determine the responsible mechanism(s) is continuing.

The hybrid codes can provide very reasonable predictions of the steady-state plasma parameters in the thrusters. For example, Fig. 7-14 shows the average profiles (along the channel axis) predicted by a 1-D model [53] for the potential, electric field, plasma density, mean electron energy, neutral density, and ionization rate for the SPT-100 Hall thruster, where 4 cm corresponds to the channel exit. The average plasma density peaks upstream of the exit, as is also predicted by the 2-D HPHall code [30] result shown in Fig. 7-15 for the SPT-100 Hall thruster channel. In both cases, there is a characteristic peak in the plasma density upstream of the channel exit in the ionization region, and a decreasing plasma density is seen moving out of the channel as the ions are accelerated in the electric field of the acceleration region.

The plasma density prediction by the 1-D code is slightly lower than the 2-D HPHall result because of differences in the heat flux calculation to the wall and the resulting values Of the electron temperature. Since the distribution function of the electrons can certainly be non-Maxwellian and anisotropic, the actual value of the density in the Hall thruster will differ somewhat from the values calculated by these existing codes.



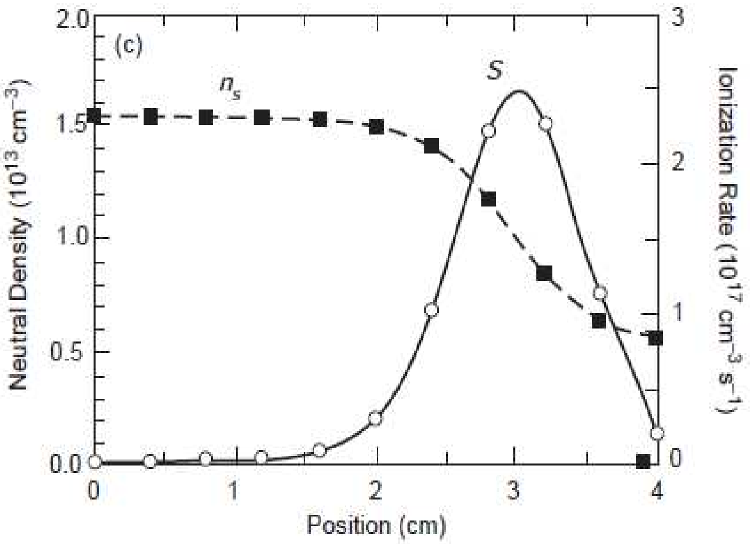


Fig. 1-14. 1-D Hall thruster code [53] for the SPT-100: (a) potential and electric field, (b) plasma density and electron energy, and (c) neutral density and ionization rate (redrawn from [53])

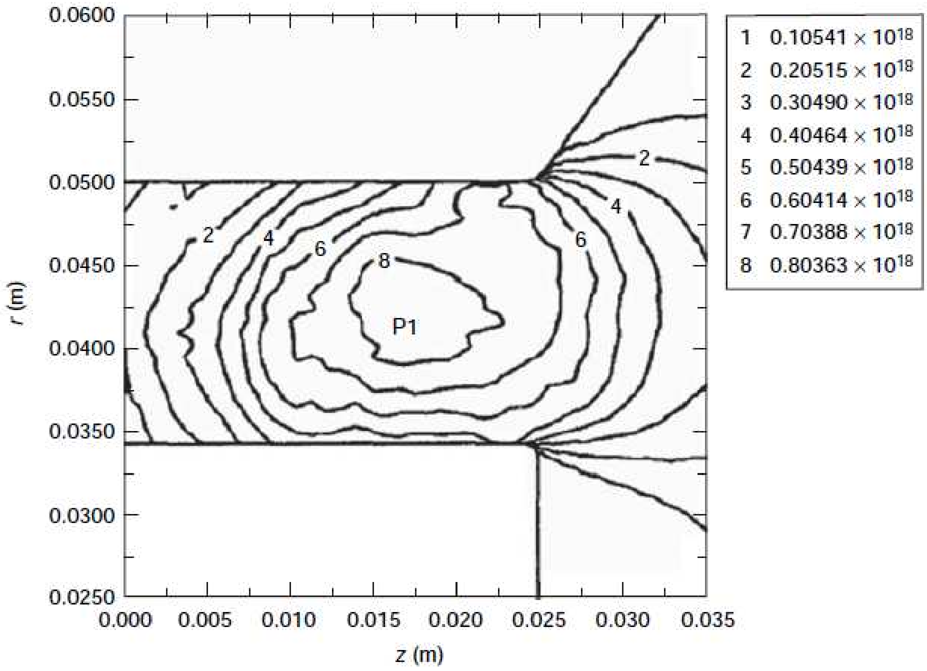


Fig. 1-15. Average plasma density computed by HPHall for the SPT-100, with the peak plasma density at P1 = 8 □□1017 m–3 (from [30])

The profiles shown by the 1-D code results in Fig. 7-14 suggest that three overlapping but distinct regions exist in the plasma channel of a well-designed Hall thruster. Near the anode, the potential drop is small due to the low magnetic field in this region, resulting in good plasma conduction to the anode but small ionization.

The ionization zone occurs upstream of the channel exit where the neutral gas density is still high and the electrons are well confined and have significant temperature. The acceleration zone exists near the channel exit where the electric field is a maximum, which occurs at this location because the magnetic field is a maximum and the transverse electron mobility is significantly reduced as described above. Outside the channel, the electric field, plasma density, and electron temperature drop as the magnetic field strength decays and the Hall current decreases.

The current versus voltage predictions from the 1-D code [53] for different values of the transverse magnetic field in the channel for the SPT-100 thruster are shown in Fig. 7-16. As the transverse magnetic field increases, the impedance of the discharge increases significantly and higher voltages are required to obtain the transverse electron mobility required to achieve the desired discharge current. Increases in the mass flow rate increase the collisional effects in the plasma region, and this results in more current at a given voltage and magnetic field. In addition, Fig. 7-16 shows regions where the 1-D code predicts oscillatory behavior, as indicted by the solid points. This is discussed in the next section

The 1-D hybrid code results shown in Fig. 7-16 suggest that the code captures the trend in the discharge impedance as the magnetic field and applied voltage are changed; i.e., the discharge current decreases as the magnetic field increases at a given discharge voltage due to the lower electron mobility.

However, the code does not predict the correct current-versus-voltage behavior for this thruster at low voltages. Figure 7-17 shows the current-versus-voltage data for one condition in the SPT-60 (a 60-mm channel outside-diameter version described in [4]). As the discharge voltage is decreased below about 200 V, the current initially increases until the energy of the electrons at very low voltage is insufficient to produce high ionization fractions, and the plasma density and discharge current then fall. Improvements in the electron transport physics are:

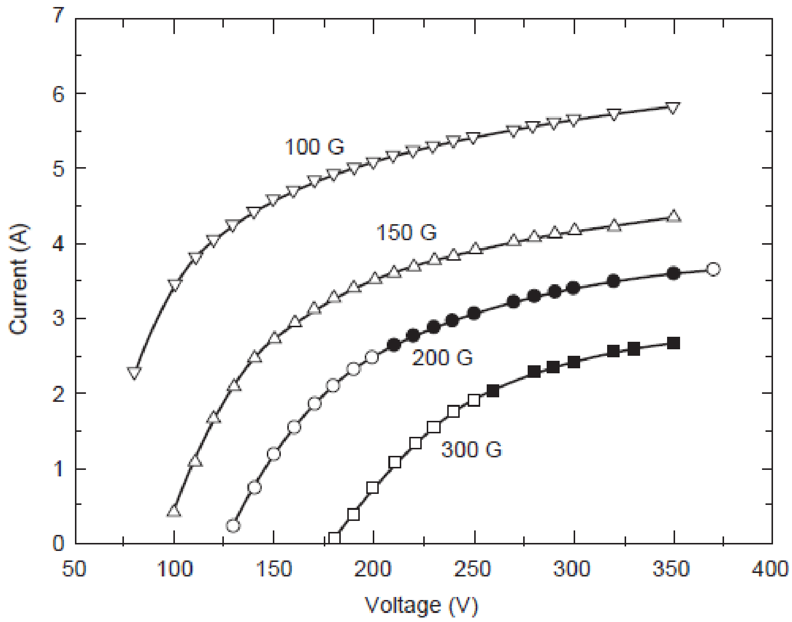


Fig. 1-16. Current-versus-voltage predictions from the 1-D code of Boeuf and Garrigues (from [53]) for the SPT-100, where the solid points indicate regions of predicted oscillations

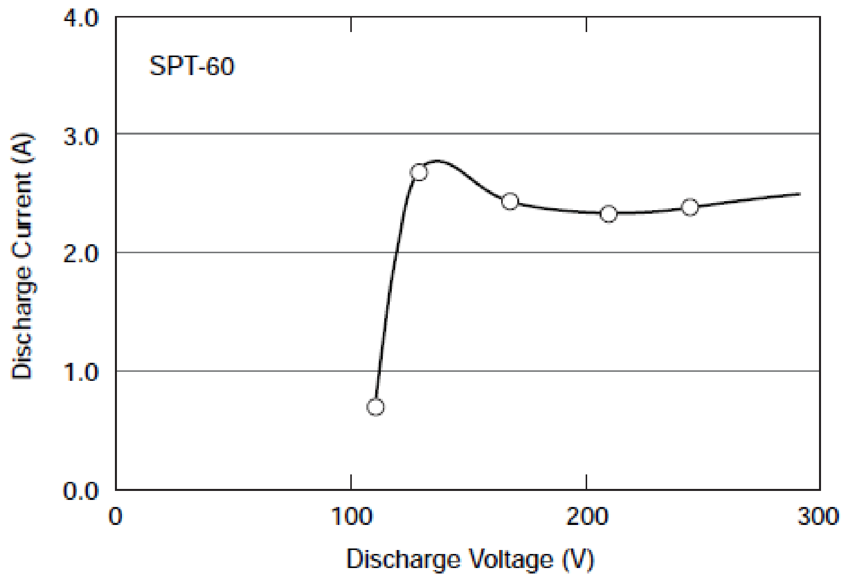


Fig. 1-17. Current versus voltage for one operating condition in the SPT-60 ([redrawn from [4]), showing the non-monotonic current variations usually observed in Hall thrusters at low discharge voltages

Clearly required for the hybrid code to fully predict the Hall thruster behavior. Work continues developing hybrid codes to better predict the thruster parameters and performance.

7.4.3 Oscillations in Hall Thrusters

Depending on their size and operating characteristics, Hall thrusters have the capability of generating many different waves and instabilities with frequencies from 1 kHz to tens of MHz. A survey of the frequencies of different plasma waves, the characteristic lengths (i.e., of sheaths, etc.) in the thruster, and wave and particle drift velocities expected in typical Hall thrusters was compiled by Choueiri [62]. The most commonly observed oscillations occur in the band of frequencies from 1–30 kHz associated with ionization instabilities and rotational oscillations in the annular discharge channel. Azimuthally propagating waves with frequencies up to 100 kHz that are not associated with ionization instabilities can also occur due to magnetic field gradients [11]. In the range of 100–500 kHz, ion transit time oscillations associated with axial motion of the ions through the ionization and acceleration regions can occur. Above this frequency range, azimuthal drift waves [63] and ion acoustic waves have also been predicted and observed.

The low-frequency time dependence of the ion and neutral behavior can be analyzed with the analytical models [30] by writing the ion conservation equation as

$$\frac{\partial n_i}{\partial t} = n_i n_o \langle \sigma_i v_e \rangle - \frac{n_i v_i}{L}, \tag{1.4.29}$$

and the neutral particle conservation equation a

$$\frac{\partial n_o}{\partial t} = -n_i n_o \langle \sigma_i v_e \rangle + \frac{n_o v_o}{L}, \tag{1.4.30}$$

where v_O is the neutral velocity and L is the axial length of the ionization zone. The perturbed behavior of the ion and neutral densities with time is linearized such that

$$\begin{aligned} n_i &= n_{i,o} + \varepsilon n_i' \\ n_o &= n_{o,o} + \varepsilon n_o', \end{aligned} \tag{1.4.31}$$

where the first term on the right-hand side denotes the unperturbed state. Combining Eqs. (7.4-29), (7.4-30), and (7.4-31) gives

$$\frac{\partial^2 n_i'}{\partial t^2} = n_{i,o} n_{o,o} n_i' \langle \sigma_i v_e \rangle^2 . \quad (1.4.32)$$

This equation represents an undamped harmonic oscillator with a frequency given by

$$f_i = \frac{1}{2\pi} \sqrt{n_{i,o} n_{o,o} \langle \sigma_i v_e \rangle^2} \approx \frac{v_i v_o}{2\pi L} . \quad (1.4\text{-}33)$$

The low-frequency oscillatory behavior of Hall thrusters is related to the velocities of the ions and neutrals relative to the scale length of the ionization zone. This indicates that periodic depletion of the neutral gas in the ionization region causes the ion density to oscillate, which impacts the electron conductivity through the transverse magnetic field and thereby the discharge current. The ionization region location can then oscillate axially in the channel on the time scale of neutral replenishment time. The models show [53] that the oscillation depends strongly on the magnetic field strength near the channel exit, and that optimum operation of the thruster generally corresponds to high mass utilization regimes where this instability occurs.

These types of oscillations, which are typically observed in the discharge current when the thrusters are operated in a voltage-regulated mode, have been called "breathing modes" [53] and "predator–prey modes" [30], and an example is shown in Fig. 7-18 for the SPT-100 Hall thruster [55]. The frequency in this experimentally observed example is about 17 kHz. However, the frequency depends on the thruster operating conditions and can range from 10 to 30 kHz for different flow rates, voltages, and magnetic fields. The 1-D numerical code However, the ionization instability driving these oscillations is the same as that analyzed in the 1-D model, and so the behavior of the instability is adequately reproduced by the 2-D model. The low-frequency oscillations can reach 100% of the discharge current depending on the voltage and mass flow (current) for a given thruster design. However, more modern designs, especially those intended for flight, typically have much lower oscillation amplitudes

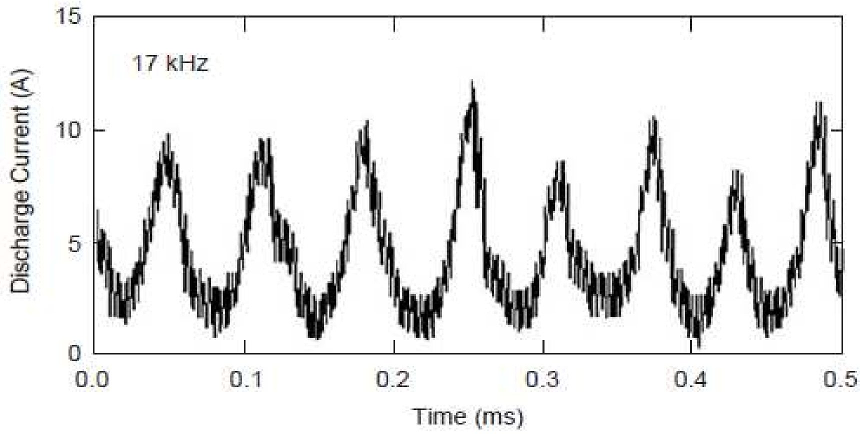


Fig. 1-18. Measured evolution of the discharge current for the SPT-100 (from [55])

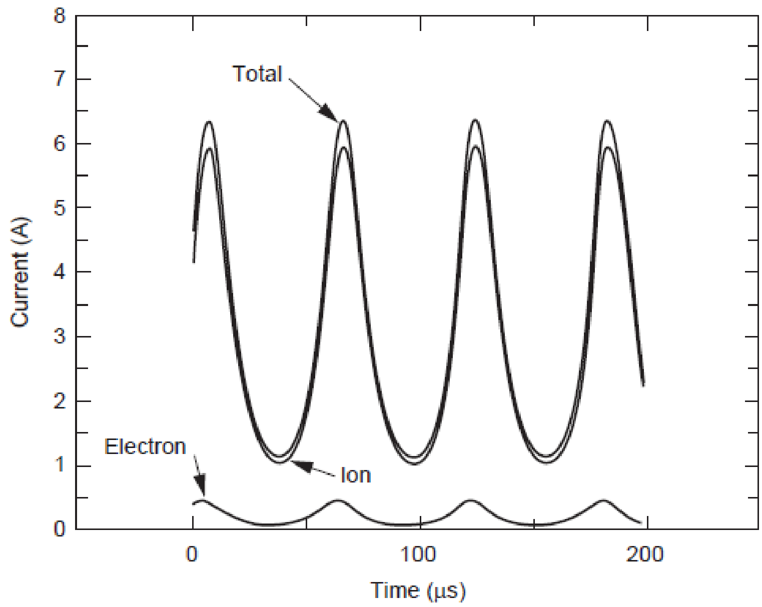


Fig. 7-19. Oscillating current predictions from the 1-D code for the SPT-100 (from [53])

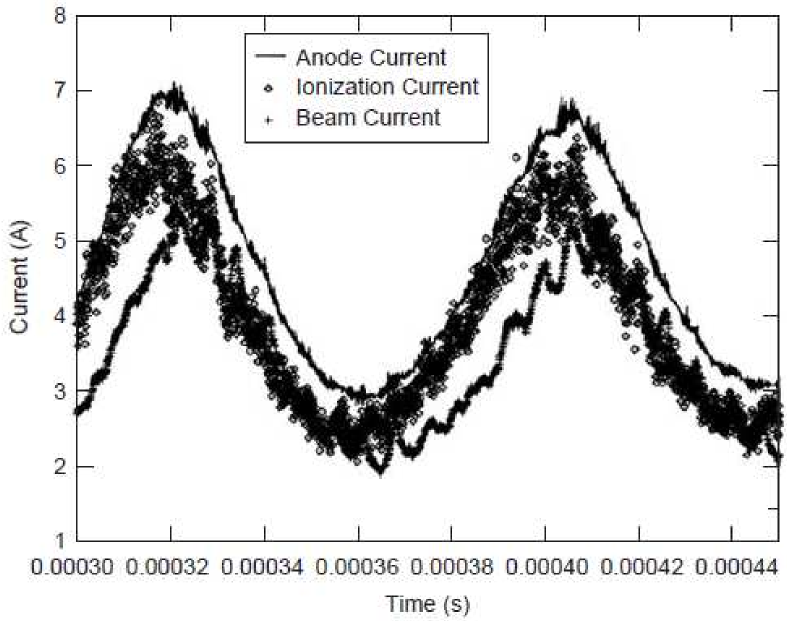


Fig. 1-20. Anode current, ionization, and beam current calculated by HPHall for the SPT-70 Hall thruster (from [30])

7.5 Hall Thruster Life

The operating time and total impulse of a Hall thruster is determined primarily by erosion of the channel wall and the life of the cathode. Hollow cathode wear-out has not represented a life limitation to date because thruster lifetimes of less than 10,000 hours are typical, and robust LaB_6 hollow cathodes have been used in all the Russian Hall thrusters. Other issues such as deposited material build-up on the electrodes, conductive-flake production, electrical shorting, etc., are also of concern in evaluating the life of a Hall thruster. However, the erosion of the channel wall by ion bombardment sputtering is a very visible process [4] that changes the channel dimensions and ultimately exposes the magnetic circuit, which, when eroded, can degrade the thruster performance. However, life tests of flight thrusters such as the SPT-100 and the PPS-1350 show that they can take hundreds to thousands of hours for magnetic circuit erosion to significantly alter thruster performance. Of greater concern, in this case, is the sputtering of iron from the magnetic circuit, which would have.

A significantly higher impact if deposited on most spacecraft components. Therefore, understanding the wall erosion rate and its dependence on thruster materials and operating parameters is of importance in predicting the thruster life and performance over time and its potential impact on the spacecraft.

The erosion rate, given by the rate of change of the wall thickness, *w*, is

$$\Re = \frac{\partial w}{\partial t} = \frac{J_i W}{\rho e A_v} Y(\varepsilon_i), \tag{1.5.1}$$

where J_i is the ion flux, W is the atomic weight, is the material density, e is the ion charge, A_v is Avogadro's number, and Y is the sputtering yield of the material, which is dependent on the ion type and energy $_i$. Since the material properties are known, the issue becomes one of knowing the ion flux, ion energy, and sputtering yield of the wall.

Several analytical models of the Hall thruster have been developed and applied to this problem [37,60,64]. The most accurate predictions have been achieved using a modified 2-D HPHall code [58] to obtain the ion fluxes and energies. The sputtering yield of boron nitride compounds used in dielectric-wall Hall thrusters has been measured by Garnier [65] versus incidence angle and ion energy, and is used in several of these models. However, the Garnier data are at only a few energies and more than 300 V. Gamero extrapolated these data to lower energies using the semi-empirical sputtering law scaling of Yamamura and Tawara [66], obtaining the following expression for the sputtering yield in units of mm^3/coulomb:

$$Y = \left(0.0099 + \alpha^2 6.04 \times 10^{-6} - \alpha^3 4.75 \times 10^{-8}\right)\sqrt{\varepsilon_i}\left(1 - \sqrt{\frac{58.6}{\varepsilon_i}}\right)^{2.5}, \tag{1.5.2}$$

Where is the incident angle of the ion? In Eq. (7.5-2), the value 58.6 represents the estimated threshold energy for sputtering required by Yamamura's model. Figure 7-21 shows an example of the yield predicted by Eq. (7.5-2) for two different incidence angles. Equation (7.5-2) was shown [58] to accurately fit the data of Garnier and provides projections of the sputtering yield down to low ion energies predicted by HPHall deeper in the channel.

Figure 1-21 shows the predicted [58] and experimentally measured erosion profiles [67] for the SPT-100 thruster inner and outer channel walls. Good agreement with the observed channel erosion is seen near the thruster exit, and the profiles have the correct functional shape. It is likely that inaccuracies in the extrapolated sputtering yield at low energies caused the disagreement with the data deep in the channel. This can be remedied by additional sputter-yield measurements at low energy and a refinement of the sputtering yield in Eq. (1.5-2).

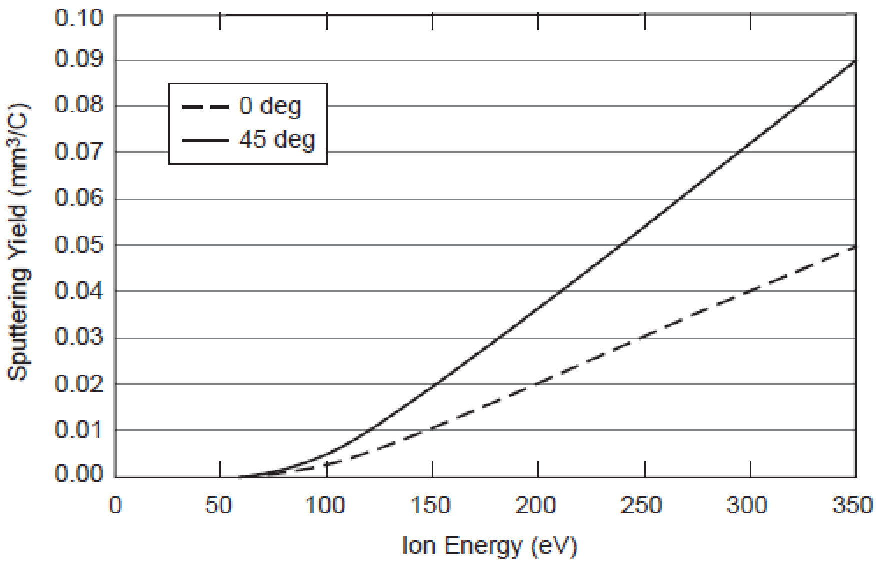


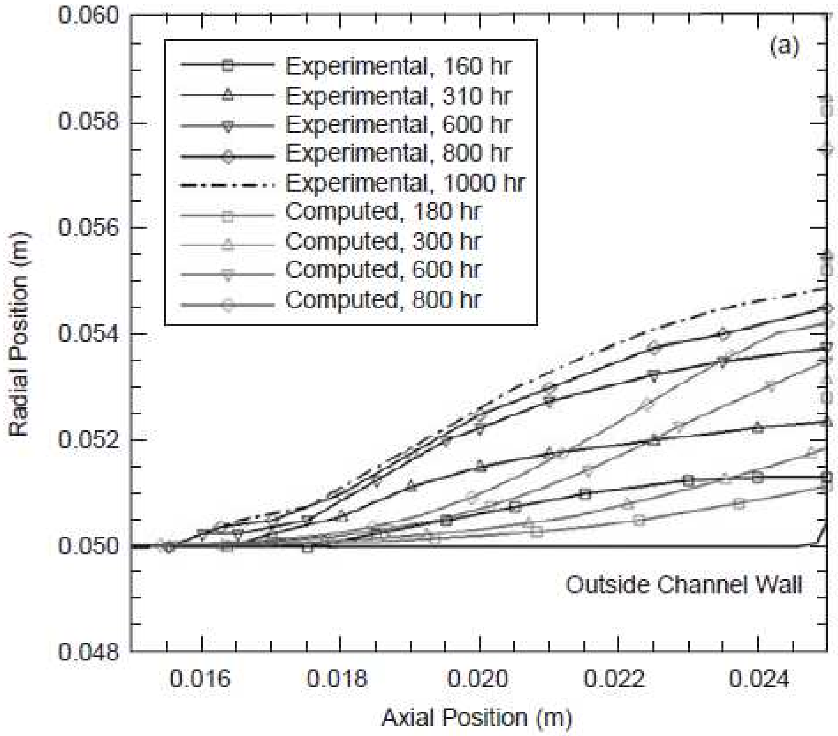
Fig. 1-21. Sputtering yield calculated for singly ionized xenon on BNSiO2 versus ion energy for two incidence angles

It is possible to develop some simple scaling rules for Hall thruster erosion in the magnetized plasma region near the exit plane. It was estimated in Section 7.3.4 that the ion flux to the wall in dielectric-wall Hall thrusters was about 10% of the beam current. It can be assumed that the energy of the ion flux to the wall is related to the beam energy, which is proportional to the discharge voltage. An examination of Fig. 7-21 shows that the sputtering yield is essentially a linear function of the ion energy. The erosion rate in Eq. (1.5-1) then becomes

$$\dot{\Re} \propto K \frac{I_b}{A_w} V_d = K \frac{I_d V_d}{\eta_b A_w}, \tag{1.5.2}$$

where K is a constant, A_W is the wall area, and Eq. (7.3-10) has been used for the beam current efficiency. Equation (7.5-3) shows that the erosion rate of the thruster wall is proportional to the power density in the accelerator channel [4].

This indicates that larger Hall thrusters are required to increase the power for a given operation time as determined by the allowable erosion of the insulator wall thickness. A good rule-of-thumb for the relationship of operation time over a reasonable throttle range of a given Hall thruster design is power * operation time = constant.



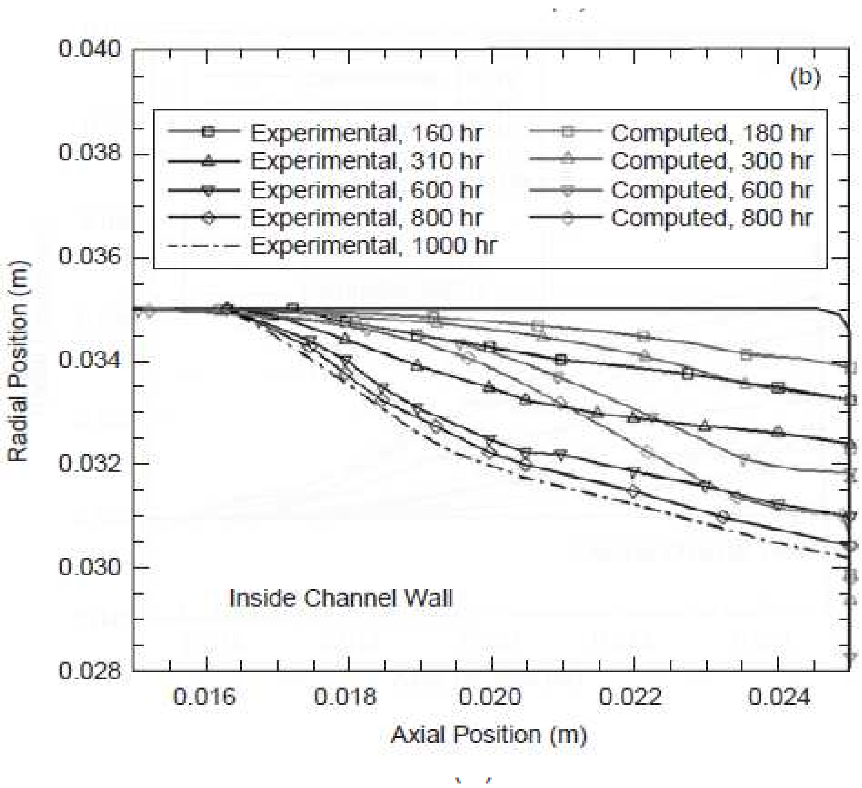


Fig. 1-22. Erosion pattern predicted by the modified HPHall code and measured for the SPT-100 thruster (redrawn from [58])

Over a limited range, the thrust from a Hall thruster is proportional to the discharge power, and so thrust * operation time = constant.

This suggests that the total impulse is essentially a constant for a given thruster design. Therefore, operating at lower thrust in throttled mission profiles will result in longer thruster operation time. However, if the throttling is too deep, the thruster performance will degrade (requiring higher input power to produce a given thrust) and the above relationship is no longer valid. Hall thruster throttle ranges of over 10:1 have been demonstrated with good performance, depending on the thruster design.

Finally, the life of TAL thrusters has not been as extensively investigated as the Russian SPT thrusters. The erosion of the channel guard rings has been identified as the primary life-limiting mechanism [29], and alternative materials were suggested to extend the thruster life by reducing the sputtering yield. Since the wall/guard ring is biased at cathode potential, the incident ion energy along the wall depends on the potential profile in the thruster channel and past the exit plane. This certainly influenced the selection of the TAL anode placement and the design of the anode/channel region to minimize the ion energy (and flux) to the walls. The dielectric-wall Hall thrusters limited the ion energy to the floating potential ($6T_e$ for xenon) for wall materials with very low secondary electron yield, and to lower energies with materials that have secondary electron yields approaching or exceeding one at the electron temperatures of typical operation. The sheath potential at the wall is likely on the order of $3T_e$ $0.3V_d$ due to space charge and non-Maxwellian electron distribution function effects. However, the lower sheath potential at the wall increases the electron flux, which results in increased power loading at the wall.

The wall material selection, therefore, is a tradeoff between efficiency and life. Dielectric walls reduce the bombarding ion energy of the wall at the expense of higher electron fluxes and higher power loading. Metallic-wall Hall thrusters have higher ion energies to the wall and therefore sputter-erosion life issues, and so they have to compensate with geometry changes to obtain the desired life. This results in higher heat fluxes to the anode, which dominates the TAL efficiency. An increase in the power of both types of thrusters also requires increases in the thruster size to obtain the same or longer lifetimes. Therefore, Hall thruster design, like ion thruster design, is a tradeoff between performance and life.

References

[1] S. D. Grishin and L. V. Leskov, *Electrical Rocket Engines of Space Vehicles*, Moscow, Russia: Mashinostroyeniye Publishing House (published in Russian) , 1989.

[2] H. R. Kaufman, "Technology of Closed-Drift Thrusters," *AIAA Journal*,vol. 23, no. 1, pp. 78–87, 1985.

[3] V. V. Zhurin, H. R. Kaufman, and R. S. Robinson, "Physics of Closed Drift Thrusters," *Plasma Sources Science and Technology*, vol. 8, no. 1,pp. R1–R20, 1999.

[4] V. Kim, "Main Physical Features and Processes Determining the Performance of Stationary Plasma Thrusters," *Journal of Propulsion and Power*, vol. 14, no. 5, pp. 736–743, 1998.

[5] A. I. Morosov and V. V. Savelyev, "Fundamentals of Stationary Plasma Thruster Theory," in *Reviews of Plasma Physics*, edited by B. B. Kadomtsev and V. D. Shafranov, vol. 21, New York: Kluwer Academic/Plenum Publishers, 2000.

[6] J. R. Brophy, J. W. Barnett, J. M. Sankovic, and D. A. Barnhart, "Performance of the Stationary Plasma Thruster: SPT-100," AIAA-92-3155, 28th Joint Propulsion Conference, Nashville, Tennessee, July 6–8, 1992.

[7] B. A. Arhipov, L. Z. Krochak, S. S. Kudriavcev, V. M. Murashko, and T. Randolph, "Investigation of the Stationary Plasma Thruster (SPT-100) Characteristics and Thermal Maps at the Raised Discharge Power," AIAA-98-3791, 34th Joint Propulsion Conference, Cleveland, Ohio, July 13–15, 1998.

[8] L. B. King, "A Re-examination of Electron Motion in Hall Thruster Fields," IEPC-2005-258, 29th International Electric Propulsion Conference, Princeton University, New Jersey, October 31–November 4,2005.

[9] R. R. Hofer, R. S. Jankovsky, and A. D. Gallimore, "High-Specific Impulse Hall Thrusters, Part 1: Influence of Current Density and Magnetic Field," *Journal of Propulsion and Power*, vol. 22, no. 4, pp. 721–731, 2006.

[10] R. R. Hofer and A. D. Gallimore, "High-Specific Impulse Hall Thrusters, Part 2: Efficiency Analysis," *Journal of Propulsion and Power*, vol. 22, no. 4, pp. 732–740, 2006.

[11] R. R. Hofer, "Development and Characterization of High-Efficiency, High-Specific Impulse Xenon Hall Thrusters," Ph.D. Dissertation, Aerospace Engineering, University of Michigan, Ann Arbor, Michigan,

[12] A. I. Morozov, Y. V. Esipchuk, A. M. Kapulkin, V. A. Nevrovskii, and V. A. Smirnov, "Effect of the Magnetic Field on a Closed-Electron-Drift Accelerator," *Soviet Physics Technical Physics*, vol. 17, no. 3, 1972.

[13] Y. Raitses, J. Ashkenazy, and M. Guelman, "Propellant Utilization in Hall Thrusters," *Journal of Propulsion and Power*, vol. 14, no. 2, pp. 247–253, 1998.

[14] L. A. Dorf, Y. F. Raitses, A. N. Smirnov, and N. J. Fisch, "Anode Fall Formation in a Hall Thruster," AIAA-2004-3779, 40th Joint Propulsion Conference, Ft. Lauderdale, Florida, July 11–14, 2004.

[15] L. Dorf, Y. Raitses, and N. J. Fisch, "Experimental Studies of Anode Sheath Phenomena in a Hall Thruster Discharge," *Journal of Applied Physics*, article 103309, part 1, vol. 97, no. 10, 2005.

[16] L. Dorf, Y. Raitses, and N. J. Fisch, "Effect of Magnetic Field Profile on the Anode Fall in a Hall-Effect Thruster Discharge," *Physics of Plasmas*, article 057104, vol. 13, no. 5, 2006.

[17] R. R. Hofer, P. Y. Peterson, A. D. Gallimore, and R. S. Jankovsky, "A High Specific Impulse Two-Stage Hall Thruster with Plasma Lens Focusing," IEPC-2001-036, 27th International Electric Propulsion Conference, Pasadena, California, October 14–19, 2001.

[18] J. A. Linnell and A. D. Galimore, "Krypton Performance Optimization in High-Voltage Hall Thrusters," *Journal of Propulsion and Power*, vol. 22, no. 4, pp. 921–925, 2006.

[19] J. A. Linnell and A. D. Galimore, "Efficiency Analysis of a Hall Thruster Operating with Krypton and Xenon Thrusters," *Journal of Propulsion and Power*, vol. 22, no. 6, pp. 1402–1412, 2006.

[20] Y. Daren, D. Yongjie, and Z. Shi, “Improvement on the Scaling Theory of the Stationary Plasma Thruster,” *Journal of Propulsion and Power*, vol. 21, no. 1, pp. 139–143,2005.

[21] V. Kim, G. Popov, B. Arkhipov, V. Murashko, O. Gorshkov, A. Koroteyev, V. Garkusha, A. Semenkin, and S. Tverdokhlebov, “Electric Propulsion Activity in Russia,” IEPC-2001-005, 27th

International Electric Propulsion Conference, Pasadena, California, October 14–19, 2001.

[22] A. I. Bugrova, N. Maslennikov, and A. I. Morozov, “Similarity Laws for the Global Properties of a Hall Accelerator,” *Journal of Technical Physics*, vol. 61, no. 6, pp. 45–51, 1991.

[23] V. Khayms and M. Martinez-Sanchez, “Design of a Miniaturized Hall Thruster for Microsatellites,” AIAA-1996-3291, 32nd Joint Propulsion Conference, Lake Buena Vista, Florida, July 1–3, 1996.

[24] O. Gorshkov, “Russian Electric Propulsion Thruster Today,” *Russian Space News*, vol. 61, no. 7, pp. 5–11, 1999.

[25] A. Smirnov, Y. Raitses, and N.J. Fisch, “Parametric Investigations of Miniaturized Cylindrical and Annular Hall Thrusters,” *Journal of Applied Physics*, vol. 92, no. 10, pp. 5673–5679, 2002.

[26] M. Keidar, A. D. Gallimore, Y. Raitses, and I. D. Boyd, “On the Potential Distribution in Hall Thrusters,” *Applied Physics Letters*, vol. 85, no.13, pp. 2481–2483, 2004.

[27] L. B. King and A. D. Gallimore, “Ion Energy Diagnostics in the Plasma Exhaust Plume of a Hall Thruster,” *Journal of Propulsion and Power*, vol. 16, no. 5, pp. 916–922, 2000.

[28] J. E. Pollard, K. D. Diamant, V. Khayms, L. Werthman, D. Q. King, and K. H. Degrees, “Ion Flux, Energy and Charge State Measurements for the BPT-4000 Hall Thruster,” AIAA-2001-3351, 37th Joint Propulsion Conference, Salt Lake City, Utah, July 8–11, 2001.

[29] C. E. Garner, J. R. Brophy, J. E. Polk, S. Semenkin, V. Garkusha, S. Tverdokhlebov, and C. Marrese, "Experimental Evaluation of Russian Anode Layer Thrusters," AIAA-1994-3010, 30th Joint Propulsion Conference, Indianapolis, Indiana, June 27–29, 1994.

[30] J. M. Fife, M. Martinez-Sanchez, and J. Szabo "A Numerical Study of Low-Frequency Discharge Oscillations in Hall Thrusters," AIAA-97- 3052, 33rd Joint Propulsion Conference, Seattle, Washington, July 6–9, 1997.

[31] E. Y. Choueiri, "Fundamental Difference Between Two Variants of Hall Thrusters: SPT and TAL," *Physics of Plasmas*, vol. 8, no. 11, pp. 5025– 5033, 2001.

[32] J. P. Bugeat and C. Koppel, "Development of a Second Generation of SPT," IEPC-95-035, 24th International Electric Propulsion Conference, Moscow, Russia, 1995.

[33] N. Gascon, M. Dudeck, and S. Barral, "Wall Material Effects in Stationary Plasma Thrusters: 1. Parametric Studies of an SPT-100," *Physics of Plasmas*, vol. 10, pp. 4123–4136, 2003.

[34] J. M. Fife and M. Martinez-Sanchez, "Comparison of Results from a Two-dimensional Numerical SPT Model with Experiment," AIAA-96- 3197, 32nd Joint Propulsion Conference, Lake Buena Vista, Florida, July 1–3, 1996.

[35] J. M. Haas, "Low-Perturbation Interrogation of the Internal and Near- Field Plasma Structure of a Hall Thruster Using a High-speed Probe Positioning System," Ph.D. Dissertation, University of Michigan, Ann Arbor, Michigan, 2001.

[36] J. A. Linnell and A. D. Gallimore, "Internal Langmuir Probe Mapping of a Hall Thruster with Xenon and Krypton Propellant," AIAA-2006-4470, 42nd Joint Propulsion Conference, Sacramento, California, July 9–12, 2006.

[37] E. Ahedo, P. Martinez-Cerezo, and M. Martinez-Sanchez, "One-Dimensional Model of the Plasma Flow in a Hall Thruster," *Physics of Plasmas*, vol. 8, no. 6, pp. 3058–3068, 2001.

[38] E. Ahedo, "Presheath/Sheath Model with Secondary Electron Emission from Two Parallel Walls," *Physics of Plasmas*, vol. 9, pp. 4340–4347, 2002.

[39] G. D. Hobbs and J. A. Wesson, "Heat Flow Through a Langmuir Sheath in the Presence of Electron Emission," *Plasma Physics*, vol. 9, pp. 85–87, 1967.

[40] V. Baranov, Y. Nazarenko, and V. Petrosov, "The Wear of the Channel Walls in Hall Thrusters", IEPC-2001-005, 27th International Electric Propulsion Conference, Pasadena, California, October 14–19, 2001.

[41] E. Ahedo, J. M. Gallardo, and M. Martinez-Sanchez, "Effects of the Radial Plasma-Wall Interaction on the Hall Thruster Discharge," *Physics of Plasmas*, vol. 10, no. 8, pp. 3397–3409, 2003.

[42] J. M. Fife and S. Locke, "Influence of Channel Insulator Material on Hall Thruster Discharges: A Numerical Study," AIAA-2001-1137, 39th Aerospace Sciences Meeting, Reno, Nevada, January 8–11, 2001.

[43] Y. Raitses, D. Staack, M. Keidar, and N. J. Fisch, "Electron-Wall Interaction in Hall Thrusters," 057104, *Physics of Plasmas*, vol. 12, 2005.

[44] I. D. Kaganovich, Y. Raitses, D. Sydorenko, and A. Smolyakov, "Kinetic Effects in Hall Thruster Discharge," AIAA-2006-3323, 42th Joint Propulsion Conference, Sacramento, California, July 9–12, 2006.

[45] N. B. Meezan and M. A. Capelli, "Kinetic Study of Wall Collisions in a Coaxial Hall Discharge," *Physical Review E*, 036401, vol. 66, 2002.

[46] E. Ahedo and F. I. Parra, "Partial Trapping of Secondary Electron Emission in a Hall Thruster Plasma," *Physics of Plasmas*, 073503, vol. 12, no. 7, 2005.

[47] E. Ahedo and V. DePablo, "Effects of Electron Secondary Emission and Partial Thermalization on a Hall Thruster," AIAA-2006-4328, 42nd Joint Propulsion Conference, Sacramento, California, July 9–12, 2006.

[48] M. T. Domonkos, A. D. Gallimore, C. M. Marrese, and J. M. Haas, "Very-Near-Field Plume Investigation of the Anode Layer Thruster," *Journal of Propulsion and Power*, vol. 16, pp. 91–98, 2000.

[49] A. V. Semenkin, S. O. Tverdokhlebov, V. I. Garkusha, A. V. Kochergin, G. O. Chislov, B. V. Shumkin, A. V. Solodukhin, and L. E. Zakharenkov, "Operating Envelopes of Thrusters with Anode Layer," IEPC-2001-013, 27th International Electric Propulsion Conference, Pasadena, California, October 14–19, 2001.

[50] D. T. Jacobson, R. S. Jankovsky, and R. K. Rawlin, “High Voltage TAL Performance,” AIAA-2001-3777, 37th Joint Propulsion Conference, Salt Lake City, Utah, July 8–11, 2001.

Problems

1. You want to design an experimental Hall thruster to operate from 100 to 800 V and from 100 to 300 gauss. Assuming that the electron temperature is always about 10% of the discharge voltage, what are the minimum and maximum lengths of the magnetized region in the channel to have a factor of 5 margin against electron and ion orbit limits? Neglect collisions.

2. Derive Eq. (1.3-42).

3. A Hall thruster has a plasma channel with a 15-cm outer diameter and a 10-cm inner diameter. Measurements made on the thruster indicate that the xenon plasma density in the channel is 5 10^{17} ions per m^3, the electron temperature T_e is 20 eV, and the radial magnetic field B_r is 200 gauss (0.02 tesla). If the thruster is operated at a discharge voltage of 300 V,

a. What is the beam power?

b. What is the electron Larmour radius r_L ?

c. What is the electron Hall parameter $_e$?

d. If the thrust correction factor = 0.9 and the mass utilization efficiency $_m$ = 0.8 , what is the thrust and Isp?

e. What is the Hall current?

4. A xenon Hall thruster has boron nitride walls with a linearly varying secondary electron yield with a value of 0.5 at zero electron energy and 2 for an electron energy of 100 eV.

a. What is the equation for the secondary electron yield in terms of the electron energy?

b. Find the equation for the secondary electron yield for a Maxwellian distribution of electron energies [hint: use Eq. (C-5)] in terms of the electron temperature T_e .

c. What is the electron temperature at which the electron flow to the wall is space-charge limited?

d. Assuming $n_e v_B$ / n_i v_i = 0.5, what is the maximum sheath potential for non-space-charge-limited flow (T_e less than the value found in part b)?

5. Assume that all the ions in a Hall thruster are produced by the Hall current ionizing the neutral gas in the channel.

a. Neglecting the ion current to the wall as small so that all the ions produced become beam ions, what is the ratio of the Hall current to the beam current if the average electron temperature is 25 eV? (Hint: write the ion production rate in terms of the Hall current and use Appendix E for ionization and excitation collision cross sections.)

b. For a xenon ion thruster with a mean radius of 9 cm, a radial magnetic field of 150 G, and a discharge voltage of 300 V, what is the ratio of the Hall current to the beam current?

6. a. xenon Hall thruster has a channel outside diameter of 10 cm and a channel width of 3.5 cm with BNSiO2 walls. Assume a plasma density of 2 10^{17} m^{-3} and an electron temperature of 20 eV in the channel with most of, much of the plasma in contact with 1 cm of the wall axially. a. What is the electron current to the wall?

b. What is the net electron current to the wall?

c. What is the power deposited on the wall associated with this electron current?

d. What is the power deposited on the wall associated with ion current?

7. Assume that the thruster in Problem 5 has alumina walls and produces 3.5 A of beam current at 400 V with an electron temperature in the channel of 15 eV. The thruster also has a beam current utilization efficiency $_b$ = 0.5.

a. What is the power into the discharge?

b. What is the total power into the alumina walls?

c. If the electron temperature at the anode is 5 eV, the mass utilization efficiency is 90%, and the thrust correction factor = 0.9, and neglecting all other power loss channels, what is the thruster efficiency?

d. For a beam voltage utilization efficiency of 0.9, how much thrust is produced?

8. The electron current to the anode in a Hall thruster can be estimated from the perpendicular electron flux diffusing through the plasma channel.

a. Neglecting the pressure gradient terms, derive an expression for the current to the anode in terms of the collision frequency in the channel plasma.

b. For the thruster in Problem 5 with a transverse magnetic field of 150 G and an axial electric field of 3 10^4 V/m, what is the anode current if only classic electron–ion collisions are considered?

c. The effective wall collision frequency can be estimated as the electron current to the *wall times the secondary electron yield and divided by the total number of particles in the plasma* ($v_w = I_{ew} / N$, *where N is approximately the plasma density times the channel cross-sectional area times the plasma length L). Derive an expression for the anode current due to the electron-wall collisions in terms of the electron current to the wall.*

d. What is the total anode current for this thruster example, using L = 1 cm for the bulk of the plasma density?

e. If the walls are made of $BNSiO_2$, what is the anode current? Why does it depend so strongly on the wall material?

9. Calculate the power lost to the wall in a xenon TAL thruster with stainless steel walls that has a plasma density at the sheath edge of 2 10^{17} m^{-3} and an electron temperature of 20 eV. The channel has a 12-cm outside diameter, an 8-cm inside diameter, and is 0.5-cm long. Which power loss channel (ions or electrons) is larger?

10. The life of a TAL thruster is limited primarily by the ion sputtering of the metallic guard rings next to the thruster exit. Assume a TAL has a plasma density near the wall of 10^{17} m^{-3} and an electron temperature of 25 eV.

a. For stainless steel walls, what is the ion current density to the walls (the guard rings) and the sheath potential?

b. If the stainless-steel sputtering yield is about 0.1 atoms per incident ion at the sheath voltage found in (a), what is the life in hours of the TAL if 2-mm thickness of the stainless-steel guard ring material can be eroded away?

c. Assume that the wall material has been changed to graphite with a secondary electron yield of about 0.5. What is the sheath potential at the wall?

d. If the graphite sputtering yield is about 5 10^{-3} atoms per incident ion at the sheath voltage found in (c), what is the life in hours of the TAL if 1-mm thickness of the graphite guard ring material can be eroded away

Chapter 8: Power System, Telecommunications and Thermal Management

Power systems for nanosatellites

8.1 Spacecraft power systems

Spacecraft power systems are responsible for generate, store, condition, control and distribute electrical power to the whole satellite, including bus and payload equipment. It is also responsible for protecting itself from failures, originated in the power system or in any other equipment in the satellite.

There are several ways to generate electrical power in a spacecraft. According to mission characteristics and constrains, designers must choose the more suitable one. Most-used available options include:

- Primary battery systems
- Fuel cell systems
- Photovoltaic-battery (PV-battery) systems
- Solar concentrator-dynamic power systems
- Nuclear-thermoelectric systems
- Nuclear or chemical-dynamic systems

Any of the above-mentioned power systems can be schematized in Figure 1 (Patel, 2005), except for the primary battery system which does not have any rechargeable energy storage.

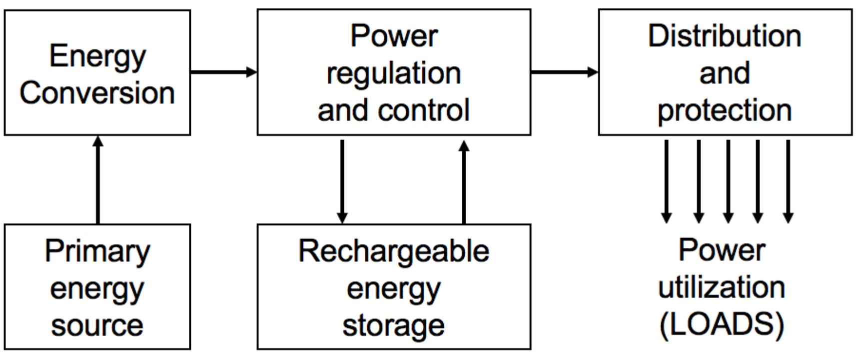


Figure 1. Basic scheme of a spacecraft power system (Patel, 2005)

This chapter is focused in PV-battery systems, which are the most used for nanosatellites. Figure 2 shows the scheme of Figure 1 applied to PV-battery based systems.

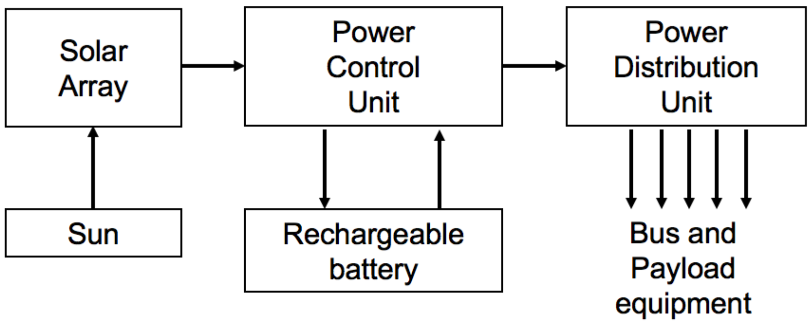


Figure 2. Basic scheme for photovoltaic-battery systems

Main working principles of PV-battery based system can be summarized as follows:

1. During sunlight periods, solar array (SA) generates enough energy to supply the loads. When solar array generates more energy than the required by loads, the battery stores this energy surplus.
2. During umbra periods, solar arrays do not generate energy, so battery supplies the energy required by the loads.
3. During peak load periods in which the energy generated by the solar array is not enough to supply the loads, battery supplies the difference.

These principles allow PV-battery systems to supply power to all the load during the whole mission lifetime.

PV-battery systems can be classified according to the way it conditions the generated power to supply the diverse loads. Selection of power system topology will be based on the loads characteristics and requirements. The three most common topologies are:

- Non-regulated bus: main bus is directly connected to the battery output. This means that bus voltage changes according to the state of charge of the battery.
- Partially-regulated bus: also, known as sun-regulated or semi-regulated bus. There is a SA regulator connected in parallel to the SA, which regulate bus voltage during sunlight periods. During umbra period, bus voltage is equal to battery output voltage.

- Fully-regulated bus: in addition to the SA regulator, there is a battery discharge regulator, thus the bus voltage is always regulated to maintain it into the required values.

Spacecraft power systems can also be divided according to its architecture. Architecture definition will depend on mission characteristics and constrains. Power system architectures are related to the way the energy is transferred:

- Direct Energy Transfer (DET): SA output power is transferred directly to the load without any regulator in series (Figure 3).

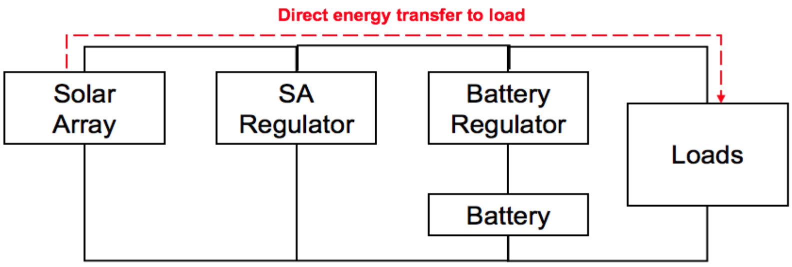


Figure 3. Direct energy transfer system

- Peak Power Tracker (PPT): SA working point (V, I) is adjusted according to the SA I-V characteristic curve and the load power consumption variation. The power point tracker device is connected in series between SA and the loads (Figure 3), generating power losses in the system.

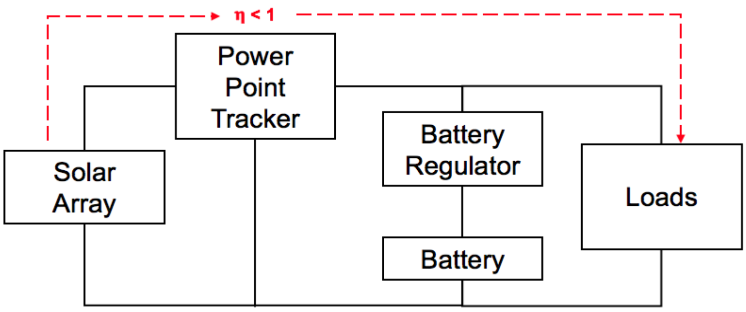


Figure 4. Power point tracker system

DET systems are more efficient, reliable and has simpler control system than PPT systems, but PPT systems allows to use smaller SA because it will be working in a point near of it maximum power point, regardless environmental conditions that normally reduce SA output power.

Nanosatellites have very limited solar array area, constrained by its reduced size, and they use body mounted or deployed solar panels without solar tracking system, which means its illumination conditions are constantly changing during one orbit. This make nanosatellites perfect candidates to use PPT systems, so the trend for nanosatellites architecture is this latest one.

Finally, spacecraft power systems can be centralized or distributed. Centralized power systems perform the voltage level conditioning itself, while distributed power systems let the voltage conversion function to each load. Main advantage of centralized architecture is the reduced number of regulators required, because a single regulator can provide power to various loads sharing the same input voltage level. This allows simpler, lighter, smaller and less expensive power systems, but they have the disadvantage of increasing the transmission losses for low voltage buses. For nanosatellites, distances are very short, so transmission losses are not a problem; instead size, mass, weight and cost are usually very limited, so centralized architecture is preferred. Voltage levels of the centralized topology are chosen to match most of, many of the loads input voltages. Nanosatellites traditionally use 3.3V, 5V, and lately 12V, for regulated buses and one unregulated bus which voltage varies according to the battery voltage.

Once defined the primary energy source, the bus regulation topology and the system architecture, power system components most be chosen according to the available technology.

8.2 Power systems components

8.2.1 Solar Array

In a PV-battery system, solar array is the primary power source. They are composed of photovoltaic solar cells which convert solar power into electric power.

Main parameters used to characterize solar cells are:

- Efficiency (η): ratio between the electrical output power of the cell and the input power received from the sun.
- I-V Curve: current Vs. Voltage characteristic (Figure 5).

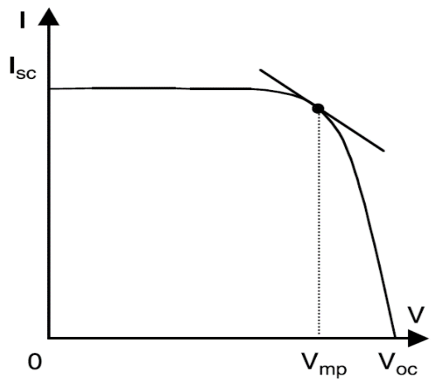


Figure 5. I-V Curve [1]

- Maximum power point (Vmp, Imp): voltage and current at which the cell generates the maximum output power.

- Short-circuit Current (Isc): current measured when shorting the output terminals.

- Open-circuit Voltage (Voc): output voltage measured when its terminals are in open circuit (no load).

I-V characteristic curve changes over the lifetime of the solar array, mainly due to three factors:

- Temperature variation.
- Solar irradiance variation (of the solar incidence angle).
- PV cell degradation

This means that the maximum output power of the solar array will change during the spacecraft mission.

Technology used for space mission are Si solar cells and Triple junction or 3J (GaInP2/GaAs/Ge), other type of cells still in research and experimental period. Nowadays most used are 3J because they have around twice the efficiency of Si cells, reaching 28%~30%. But 3J cells cost still much higher than Si ones, so few very low-cost missions, typically nanosatellites, chose to use Si cells.

Solar cells are connected in series to reach a desirable voltage, forming strings. Several strings are connected in parallel to reach the desirable output power, forming the solar array. This array need to be installed in a mechanical support to form the solar panel. Solar panels can be classified according to the construction of their mechanical support:

- Body-mounted: solar cells are attached directly on the spacecraft body surface. They have high reliability because there is not any mechanism involved, but they have no control over the solar incidence angle, making them very inefficient. This kind of solar panels are widely used in nanosatellites.
- Rigid: solar cells are mounted in rigid substrates that use hinges to allow them to be folded during launch. They can be fixed or have a solar tracking system to improve the solar incidence angle during the mission. Nanosatellites also use this kind of deployable rigid panels to increase the SA area, usually without solar tracking system.
- Flexible: utilize flexible thin-film cells mounted on a flexible structure

densely packed during launch. Once on-orbit they are deployed or rolled-out.

8.2.2 Power Control Unit

Power Control Unit or Power Conditioning Unit (PCU) is the interface between SA, battery and load buses. It main functions are to supply electrical power to the bus or buses, to regulate bus voltage according to the topology chosen, to control the battery charge and discharge processes and to communicate with on-board computer.

It is mainly composed by:

- Solar array regulator: regulates the output power of the solar array (only DET architecture with fully or partially-regulated bus topologies).
- Battery charge regulator (BCR): regulates the battery current/voltage during the charging process.
- Battery discharge regulator (BDR): regulates the bus voltage when the battery is supplying the power to the loads (only for fully-regulated bus topology).
- Sub-computer: communicates with on-board computer and performs battery charge/discharge control functions.
- Power point tracker (PPT): maintains the SA operation point to match the power load consumption (only PPT architecture).
- Filters: filter bus ripple to keep it below the required limit.

8.2.3 Power Distribution Unit

The power distribution unit (PDU) ensures that all loads, except critical and essential loads, are powered safely through switches and fuses or relays. The protection devices are not aimed to protect the loads, but to protect the power system from faults in the user equipment.

In some platforms, power control and power distribution functions are performed by the same unit, called Power Control and Distribution Unit (PCDU). Due to the limited space, available and the simplicity of the loads, nanosatellites usually implement this solution.

8.2.4 Battery

In a PV-battery system, the battery function is to provide electrical energy to loads during time periods where the solar array cannot generate enough energy to supply load requirements. These periods are umbra periods and peak load periods.

Main parameters used to characterize batteries are:

- Capacity [Ah]: total amount of electricity produced in the electrochemical reaction.
- Depth of discharge (DOD) [%]: percentage of discharged capacity (Ah) expressed as a percentage of the rated capacity.
- Energy density [Wh/unit]: amount of electrical energy delivered per unit mass (Wh/kg) or per unit volume (Wh/l).
- Cycle Life or Lifetime: number of charge and discharge cycles that the battery can support without failing to meet specific performance criteria, like capacity reduction.
- Cut-Off Voltage [V]: cell minimum allowable voltage.
- Charge Voltage [V]: battery voltage when it is full charged.

There are several battery technologies historically used in space missions, like Ni-MH, Ni-H2, Ni-Cd, Li-Ion (Lithium-Ion) and Li-Po (Lithium-Polymer), but the trend is to use the lithium bases ones, especially for nanosatellites that requires high energy density. Other than their high-energy density (about twice of Ni-Cd), Li-Ion cells have high cell voltage (equivalent to three times those of Ni-Cd), low self-discharge rate, do not have memory effect and they allow cells parallel connection to increase the battery capacity. Nevertheless, these cells are fragile and vulnerable to overload, shock that compromise their physical integrity, overheating and short circuits are some of the factors that could carry to unwanted risk situations, so protection and monitoring circuits are needed for each cell within the battery. Li-Po batteries differ from Li-Ion because it uses Polymer Gel Electrolytes (GPEs). Nowadays most batteries marked as Li-Po are actually a hybrid system which correct name is Li-Ion Polymer. This kind of battery can achieve energy density slight higher than the Li-Ion, as well as an improvement in the safety factor, being more resistant to overload and less vulnerable electrolyte leakage.

8.2.5 Harness

Harness function is to electrically connect al equipment. It includes insulated conductors, shielded wires and connectors.

References

Patel, M. R. (2005). Spacecraft Power Systems. Unite States of America: CRC Presss.

8.3 Satellite Communications

Communicating to your satellite is a vital subcomponent on any mission into space. Satellites communicate to ground stations on Earth, transmitting telemetry data, spacecraft health status, scientific experimental data, remote sensing data and images etc. The transmitted frequency from the Earth up to the spacecraft is the uplink sending commands. Once the transmission is received by the satellite the signal passes through a mixer where the frequency is then modulated from the uplink frequency to the receivers' frequency. Then passed along to interpret the data transmitted from Earth. Once the information is interpreted the signal goes through a Low Noise Amplifier (LNA), and is mixed into the transmitting frequency the signal is sent through the transmitting antenna to Earth. This is called a downlink, where the satellite transmits data down to the Earth ground station. On your spacecraft, there needs to be a full-duplex system meaning your satellite can receive and transmit data instead of just one. The uplink and downlink are very important to have different frequencies to avoid high noise figures in signal and to avoid transmission interference between the weaker uplink signal compared to the more powerful downlink signal. There are many ways to integrate antennas onto your spacecraft, the type of antenna will be determined in early mission planning to determine what communication system is optimal for your mission success.

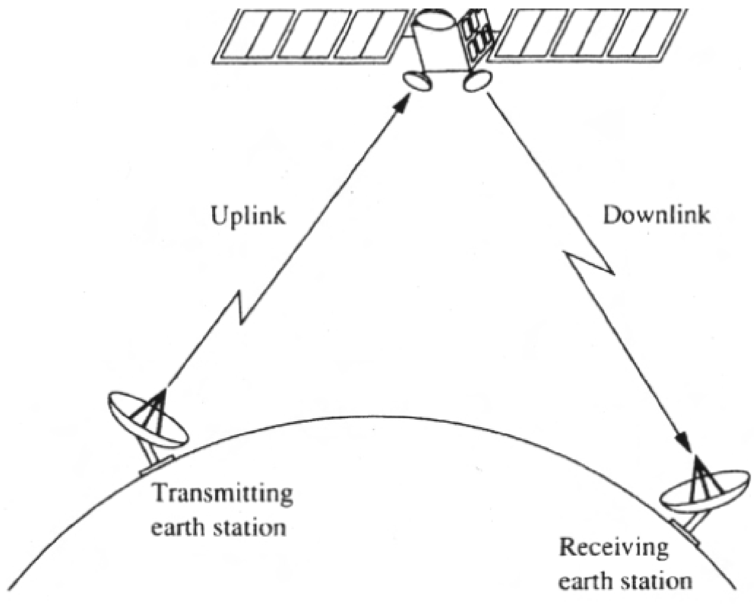


Figure 8.3.1 Satellite uplink and downlink network

8.3.1 Satellite Antenna

Satellites a required to have transmitting (downlink) and receiving (uplink) antennas. These antennas are separate so the signals and transmissions do not get interfered and to reduce noise in the system. SmallSats are usually orbiting in LEO (Low Earth Orbit) with 90 minuet orbital period. It is critical that you know the trajectory and coverage of your ground station so you will receive optimal telemetry and be able to communicate with your spacecraft for the longest period during the coverage of your ground tracking stations. The higher the gain and larger size of the dish of the antenna on the ground the better success you will have with your communication link. Satellites orbiting the Earth must have precision tuned frequencies for transmission with high power to ensure the signal once passed through the Earth’s atmosphere and receiving noise to the signal also from the ground noise on Earth effecting your ground station and downlink signal. Satellites need to have a specific EIRP (effective isotropic radiated power) to not interfere with other signals and get distorted and noisy. They also need EIRP to avoid interference with terrestrial communication systems. Weather and rain will also affect your signal as well, so the EIRP is the amount of transmitted power for the signal to downlink to Earth stations without as many complications.

Larger communication satellites have larger antenna arrays and larger antenna dish sizes to supply their mission with success. Small satellites are usually equipped with rod antennas or patch antennas to account for the mass and form factor restraints. Also, SmallSats are not designed for communications unless they are in a constellation providing global surveillance and communicating. Most SmallSats operate in the 433MHz frequency range in UHF, Ku, Ka, S, X and L band. When designing the telecommunication system for your satellite you need to figure the: wavelength of the uplink frequency, diameter of downlink antenna, diameter of uplink satellite, isotropic gain of uplink and downlinks, and gains of all antennas in the system.

8.3.2 Ground Antenna

Earth ground stations are dependent on two factors for uplink and downlink: isotropic gain to ensure maximum carrier power to noise ratio for receiving, second is the beam width being narrow enough when transmitting to spacecraft so it does not interface with other satellites. The 3-dB beam width needs to be narrow enough to not receive signals from other satellites and to prevent satellites from receiving the wrong transmission from Earth. The antenna size and gain will control these two aspects.

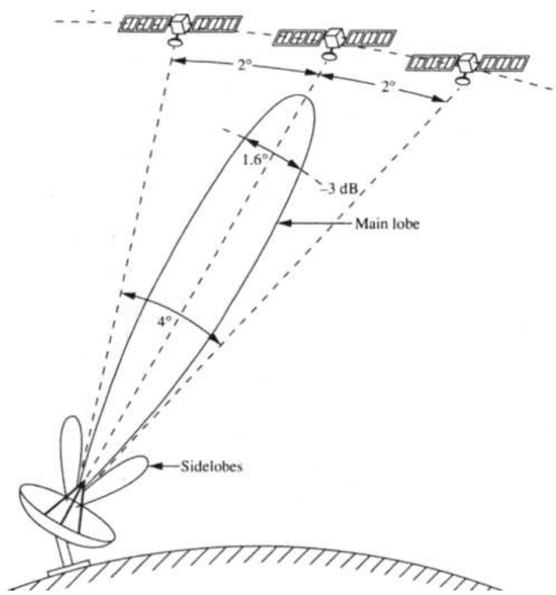


Figure 8.2.2. Frequency allocations

VHF Band – Very High Frequency

136-138MHz- Restricted.

144-146MHz- Amateur satellite band

148-150 MHz – Uplink for satellites that down link (136-138MHz)

240-270MHz- Military use

UHF Band- Ultra High Frequency

399.9-403MHz- Positioning, navigation, mobile communication.

432-438MHz- Amateur satellite band

460-470MHz- Environmental and remote sensing

L Band-

1.2 -1.8 GHz- GPS and mobile satellite communication band.

1.67-1.71GHz-Meterology satellite imaging and data link.

S Band-

2.025 – 2.3GHz- Space and deep space missions. Used for Apollo lunar missions.

2.5-2.67GHz- Point to point (fixed) communication satellites

C Band-

3.4-4.2GHz- Fixed Satellite and broadcast satellite service.

X Band-

8-9GHz- Space research and deep space missions, military and environmental, most satellites have this band on board.

Ku Band-

10.7-11.7GHz- Fixed Satellite Service

11.7-12.2GHz- Downlink for TV programs broadcasting satellite service

14.5-14.8GHz- Uplink for (11.7-12.2GHZ)

17.3-18.1GHz- Alternative Band

Ka Band-

23-27GHz- Fixed link, broadcasting, environmental and space operation satellites.

8.3.3 Link Analysis

Below are equations to help define and determine the best fit scenario for your satellite mission. They are known as the Link Analysis Equation and all are important factors that are critical for telecommunication link of any satellite.

Friss Equation

$$P_r = \frac{\lambda^2 G_t G_r P_t}{(4\pi)^2 r^2}$$

Noise Power Level

$$N_{sys} = kT_{sys}B$$

System Noise Temperature

$$T_{sys} = T_e + T_{ext} + T_{ohm}$$

variables:

T_{sys}- system temperature

T_e - effective input noise temperature of the receiver

T_{ext}- external atmospheric noise

T_{ohm}- noise from ohmic loss in antenna from transmission line

Carrier to noise ratio

$$\frac{C}{N} = \left(\frac{P_t G_t}{kB}\right)\left(\frac{\lambda}{4\pi r}\right)^2\left(\frac{G_r}{T_{sys}}\right)$$

Link analysis equation for the C/N ratio in decibels

$$\left(\frac{C}{N}\right)_{\mathrm{dB}} = [10\log(P_tG_t) - 10\log(kB)] + [10\log(G_r) - 10\log(T_{\mathrm{sys}})] - \left[20\log\left(\frac{4\pi r}{\lambda}\right)\right] - \text{other losses in dB}$$

Link analysis equation with factor expressed in decibels

$$\left(\frac{C}{N}\right)_{\mathrm{dB}} = 228.6 + \mathrm{EIRP} - 10\log(B) + G_r - 10\log(T_{\mathrm{sys}}) - L_p - L_a$$

variables;

B = bandwidth in Hz
T_{sys} = system noise temperature in K
L_p = free-space path loss in dB
L_a = atmospheric loss in dB
EIRP = P_tG_t = effective isotropic radiated power in dBW
G_r = receiving antenna gain in dB

Works Cited

Pratt, Timothy, Charles W. Bostian, and Jeremy E. Allnutt. Satellite Communications. New York: Wiley, 2003. Print.

Sellers, Jerry Jon., Wiley J. Larson, William J. Astore, Anita Shute, and Dale Gay. Understanding Space: An Introduction to Astronautics. New York: McGraw-Hill, 1994. Print.

Wertz, James Richard. Space Mission Analysis and Design. El Segundo: Microcosm, 2003. Print.

Satellite Thermal Control

Contents

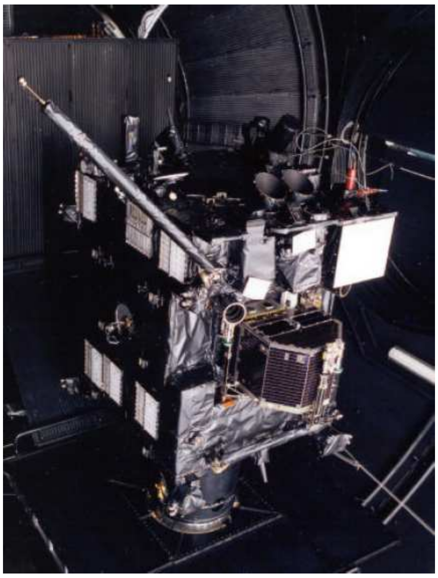
- **Heat transfer Basics**
- Conduction
- Radiation
- Importance of thermo-optical properties
- **Role**
- Why thermal control required?
- **Design**
- What is thermal Design?
- What types of S/C design exist?
- **Means**
- What to control the flux/temperatures?

Heat transfer basics

- Heat transfer Basics

1. Satellite Heat transfer modes
2. Conduction
3. Radiation

ROSSETTA* FM



* Without solar panels

Conduction

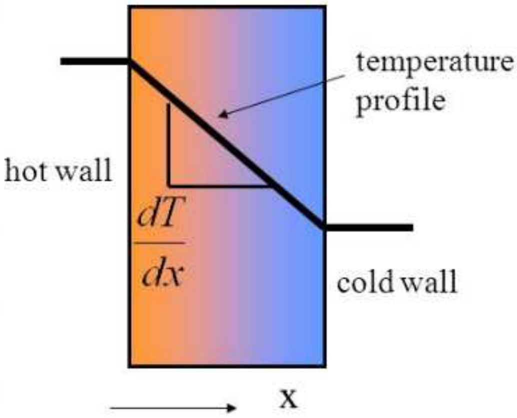
- Definition
- Propagation of energy from particle to particle
- In solid, liquid or gaseous continuous matter, homogeneous or not
- Without matter displacement

Fourier's Law

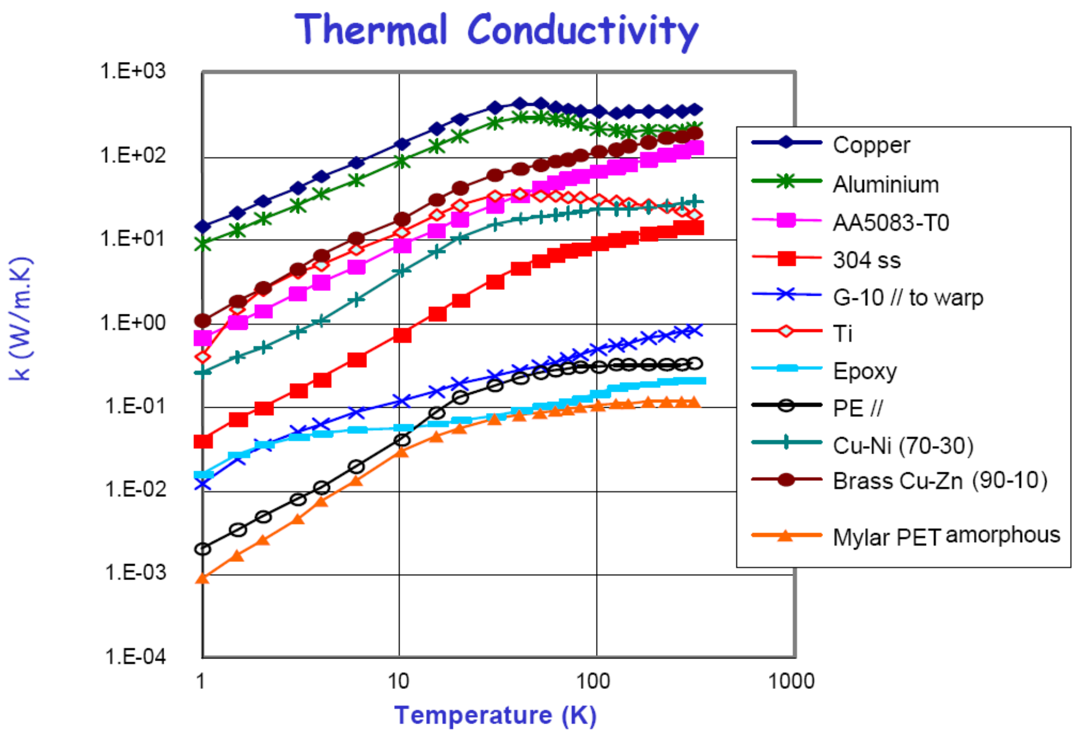
- The heat flux is proportional to the temperature gradient:
- $$\frac{Q}{A} = q = -k\nabla T \qquad \text{...... (1)}$$

Where k=thermal conductivity (W/m°C or Btu/h ft °F)
-- a measure of how fast heat flows through a material
-- k(T), but we usually use the value at the average temperature

q can have x, y, and z components; it's a vector quantity



$$\dot{Q} = -k \cdot A \frac{dT}{dx} \qquad \text{..... (2) } \textbf{\textit{Fourier's Law}}$$



Radiation

- Characteristics
- Propagation of electro-magnetic energy in straight line
- Between surfaces separated by
- Absorbing, scattering media
- Or in vacuum

Hence without matter displacement and reflected, absorbed or transmitter on surrounding bodies.

Source:

Thermal agitation of particles.

Black body Radiation

Black Body- is real or fictitious surface that absorbs all incident radiant energy i.e. from every direction at every wavelength.

$$\boxed{\alpha(\theta,\lambda)=\alpha=1}$$

Isotropic emitter radiated energy depends only on temperature.

$$\boxed{\varepsilon(\theta,\lambda)=\varepsilon=1}$$

- Black Body Emitted Energy

hemispherical spectral

hemispherical total emissive

emissive power

emissive power

Planck's Law

$$E_{\lambda,T} = \frac{2\pi\, h c^2}{\lambda^5 \left(e^{\frac{hc}{k_B \lambda T}} - 1 \right)}$$

Stefan Boltzmann Law

$$E_{bb,T} = \int_0^\infty E_{\lambda,T}\, d\lambda = \sigma T^4$$

Black body Radiation

- Radiated Energy between Black Bodies

$$Q_{ij} = A_i\, F_{ij}\, \sigma \left(T_i^4 - T_j^4 \right)$$

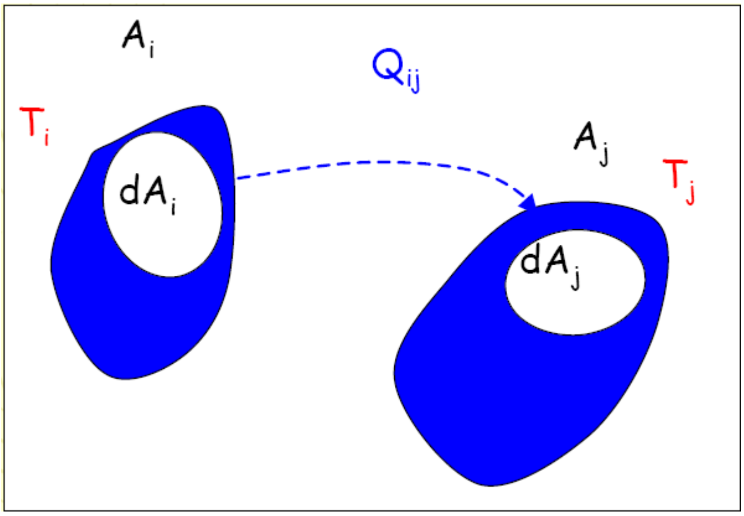
with F_{ij}, the view factor between surface i and surface j or:

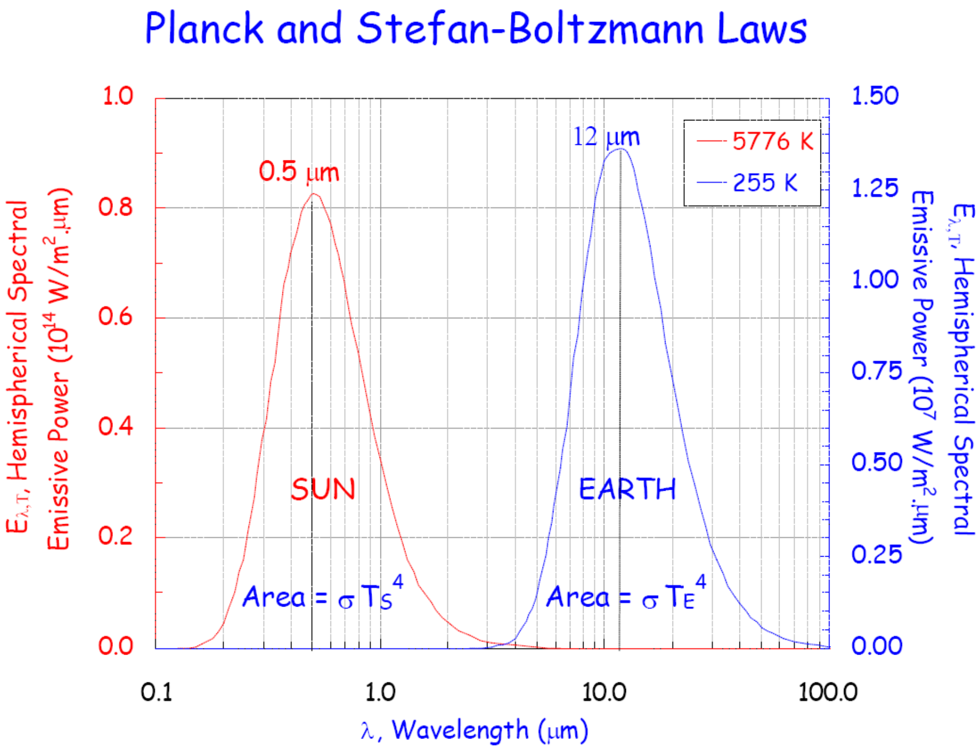
$$Q_{ij} = A_i\, \sigma \left(T_i^4 - T_j^4 \right)$$

when:

$$F_{ij} = 1$$

Visual of Blackbody energy radiated between Black Bodies.





Radiation- real body can absorb, reflect or transmit radiation energy.

-- all parameters are wavelength and angular dependent

-- general case: semi-transparent

$$\alpha(\lambda)+\rho(\lambda)+\tau(\lambda)=1$$

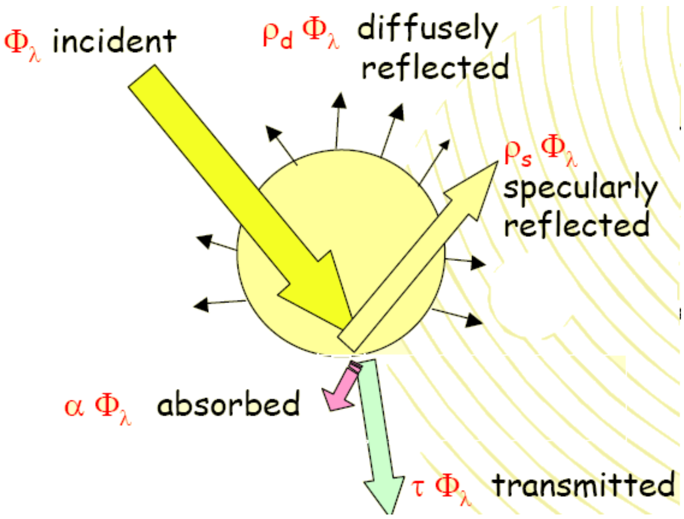
$$\rho=\rho_s+\rho_d$$

--opaque

$$\tau(\lambda)=0$$

hence

$$\alpha(\lambda)+\rho(\lambda)=1$$



Radiation- real body

Surface Emissivity

- ratio of surface radiated energy to that of a black body at the same T
- always <1 for a real surface for a black body

$$\varepsilon(\theta,\lambda)=\frac{\int_0^\infty \alpha_{\lambda,T}\,E_{\lambda,T}\,d\lambda}{\int_0^\infty E_{\lambda,T}\,d\lambda}<1 \qquad \varepsilon(\theta,\lambda)=\varepsilon=1$$

For black body

- depends on direction q and wavelengthλ of emitted energy

– therefore can be

-- directional (d) or hemispherical (h)

-- spectral (s) or total (t)

-- averaged over all directions, wavelengths or both

Radiation- real body

Surface Absorptivity

- ratio of surface absorbed energy to incident energy
- always <1 for a real surface

For a black body

$$\alpha(\theta,\lambda)=\frac{\int_0^\infty \alpha_{\lambda,T}\,E_{\lambda,T}\,d\lambda}{\int_0^\infty E_{\lambda,T}\,d\lambda}<1 \qquad \alpha(\theta,\lambda)=\alpha=1$$

- depends on incident energy direction q and wavelength λ

– therefore can be

- directional (d) or hemispherical (h)
- spectral (s) or total (t)
- averaged over all directions, wavelengths or both

Radiation- real body

Absorptivity vs Emissivity

– for a given direction q and at any wavelength λ

2nd Kirchhoff's Law

$$\boxed{\alpha(\theta,\lambda)=\varepsilon(\theta,\lambda) \quad \forall\theta,\forall\lambda}$$

– in general, hemispherical total values are different

$$\boxed{\alpha \neq \varepsilon}$$

because, α and ε have a strong wavelength dependence

– source temperature of incident radiation (Sun at 5776 K) different than surface temperature (satellite –250 -> 300°C)

Radiation- real body

- Solar Absorptivity α_S and Hemispherical Emissivity ε_H
- α_S is the solar absorptivity refers to UV wavelengths
- $\alpha_S = \varepsilon_S$ integrated over 0.2-2.8 mm i.e. 95% solar spectrum
- ε_H is the hemispherical emissivity refers to IR wavelengths
- $\alpha_H = \varepsilon_H$ integrated over 5-50 mm i.e. body at –250/300°C

  but $\alpha_S \neq \varepsilon_H$ because the spectra are different.

Radiation- data

Spectral Reflectance *(ρ)*

– Zinc Oxide Potassium

– Silicate Coating

Black Body Emittance

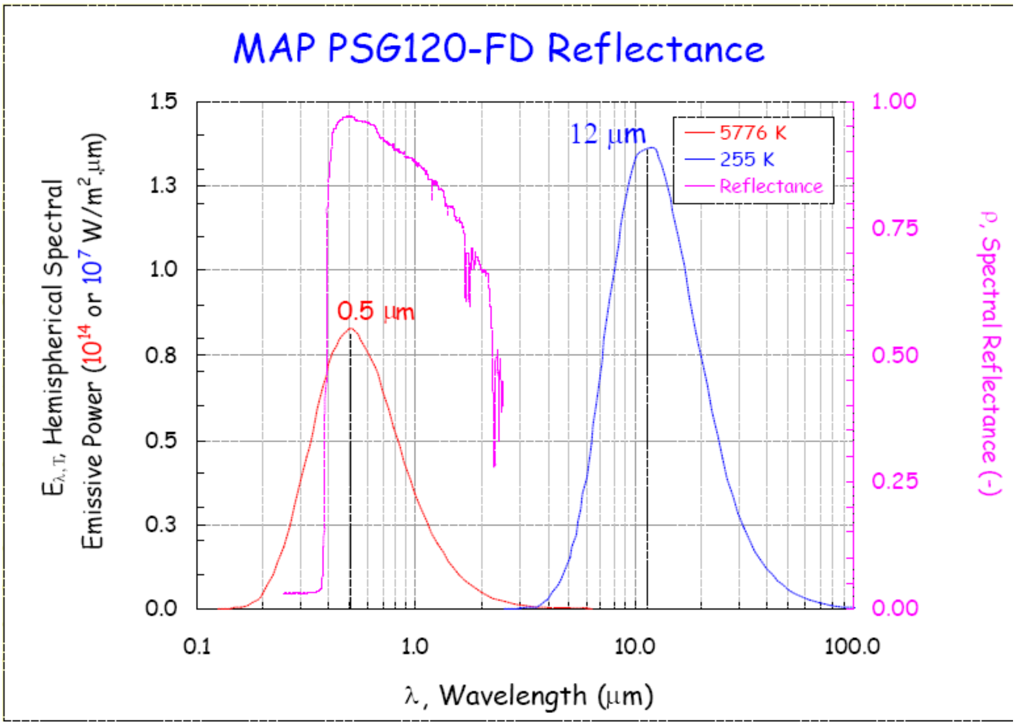
– integration over solar

– (5776K) wavelengths

– α_S =0.20

- integration over infrared black body (300K) wavelengths

ε_H =0.87



Radiation- data

Finish	α_S	ε_H	α_S/ε_H
VD Au	0.23	0.03	9.20
VDA	0.15	0.05	3.00
black paint	0.94	0.81	1.16
white paint	0.20	0.88	0.23
SSM (Ag 2 mils)	0.10	0.60	0.17
OSR	0.09	0.82	0.11

Role

maintain within Specified Ranges

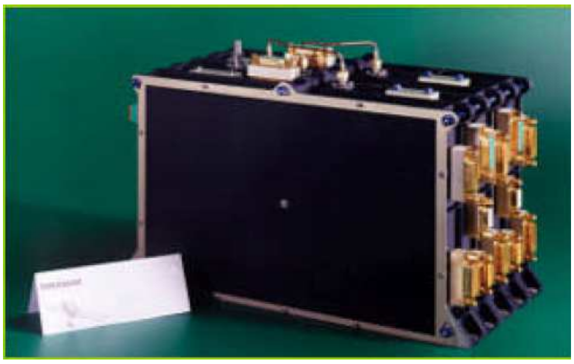
– temperatures

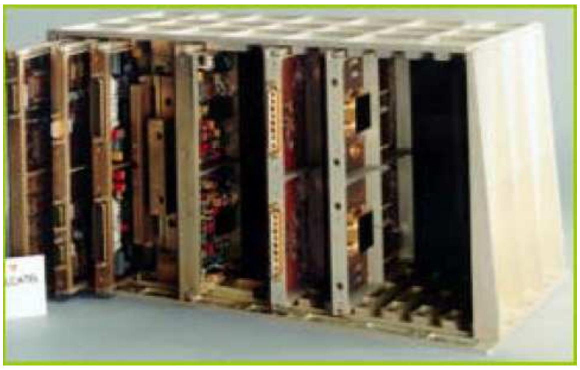
– temperature gradients (K/length)

– temperature stability (K/time)

– radiative/conductive heat flow (W)

On Board Data Compression Unit (ALS) SPOT 5 Solid State Recorder (ALS)





Role

- of What?
- electronic units
- instrument e.g. optical bench
- S/C structure
- interface between modules

Visual Monitoring Camera

SOHO, STM

PL

SV

Role- key requirements

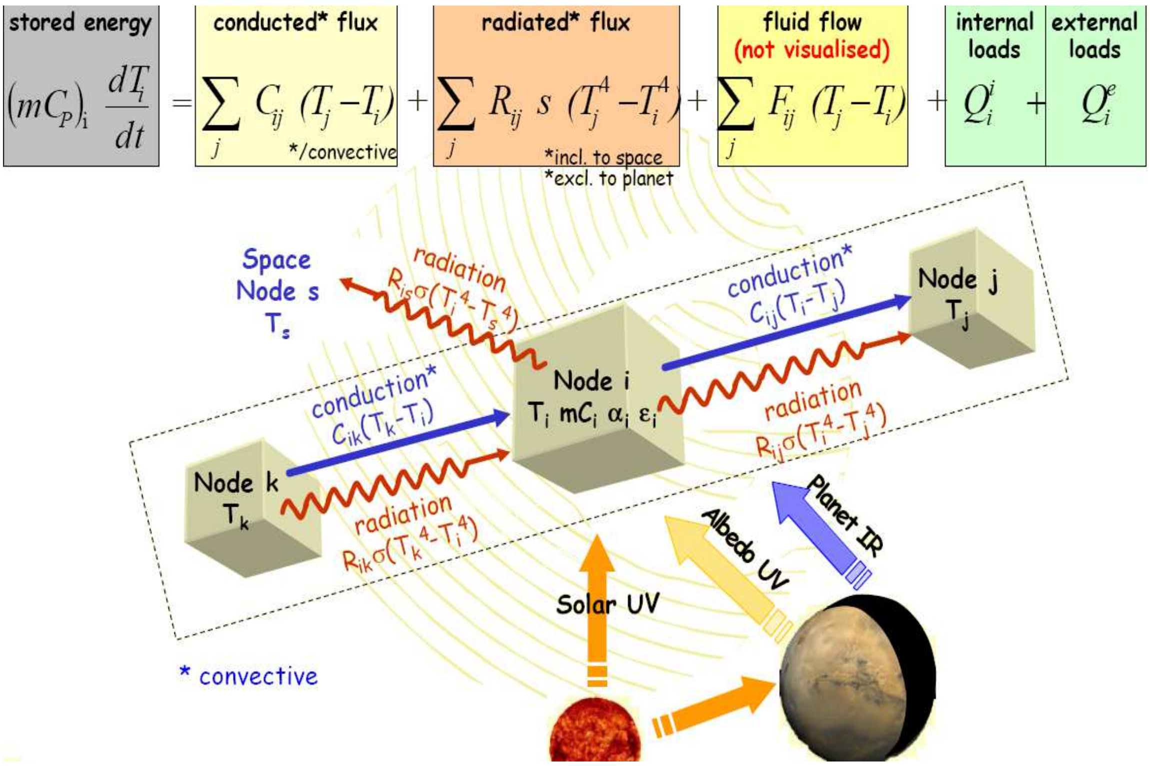
- Narrow Temperature Ranges
- – electronics equipment
- classical equipment [-10, +40] °C
- battery [0, +20] °C
- propulsion system [+10, +50] °C
- limited Temperature Gradients
- DT < 5°C across optical instrument (1.5 m)
- DT/Dx < 2°C/m for structural element

DT < 5°C between MMH and NTO tanks

- Stable Temperatures
- DT/Dt < 5 K/h for typical electronic unit
- DT/Dt < 0.1 K/mn for CCD camera
- DT/Dt < 100 µK/mn for cryogenic telescope

- Why is it so important?
- – low temperatures for reliability of components
- – narrow temperature ranges for sensitivity of detectors, units
- – small temperature gradients for pointing of instruments, S/C

Design



Design

- Balance HEAT FLOWS

to fulfil

- REQUIREMENTS

results in

- TEMPERATURES
- through Heating

– Absorb from external sources (solar, albedo, planet IR)

- selective coatings
- use the internal sources
- electronic dissipations, MLI insulation efficiency

– dissipate heat internally

- heater
- RTG, RHU

Design

- **transfer heat from hot area**
- conduction, radiation
- latent heat of evaporation/condensation
- **or through Cooling**

– reject to deep space (3 K)

- with low a/e radiative coatings on radiators

– transfer heat to cold area

- by conduction,
- radiation
- through condensation/boiling in fluid loops or heat pipes

– through cryogenic techniques

- cryostats
- coolers (Peltier, Joule-Thomson…)

– ablation

Design-Radiative concept

- Principle

– when internal power dissipation small w.r.t. external absorbed energy

– balance between

- absorbed incident radiant energy (solar…)
- emitted radiant energy (sT4)
- Characteristics

– no insulation

– average temperature driven by

• external fluxes

– local temperature hot spots still possible

PROBA1

Design- Insulated concept

• Advantages w.r.t. Radiative Concept

– less sensitive to

• eclipses

• external loads changes

– temperatures are more uniform

– little ageing of un-irradiated coatings

We are between those 2 CONCEPTS

ASTRA-1K FM Antenna Deployment

Means – limitations

- **Cooling Limitations**

  – radiator 100 mW at 100 K for 0 W dissipation

  – cryo-coolers 1 W at 50 K for 100 W dissipation

  – liquid He few mW at 4 K for 1 ton/2 years

- **Heat Transport Limitation/Performance**

  – conduction (pure Al tube k=200 W/m.K)

- 1.5 W @20°C l=1.00 m Ø=2 cm DT= 25 K m= 0.8 kg
- 11 kW@20°C l=0.70 m Ø=4.04 m DT= 3 K m= 24 t

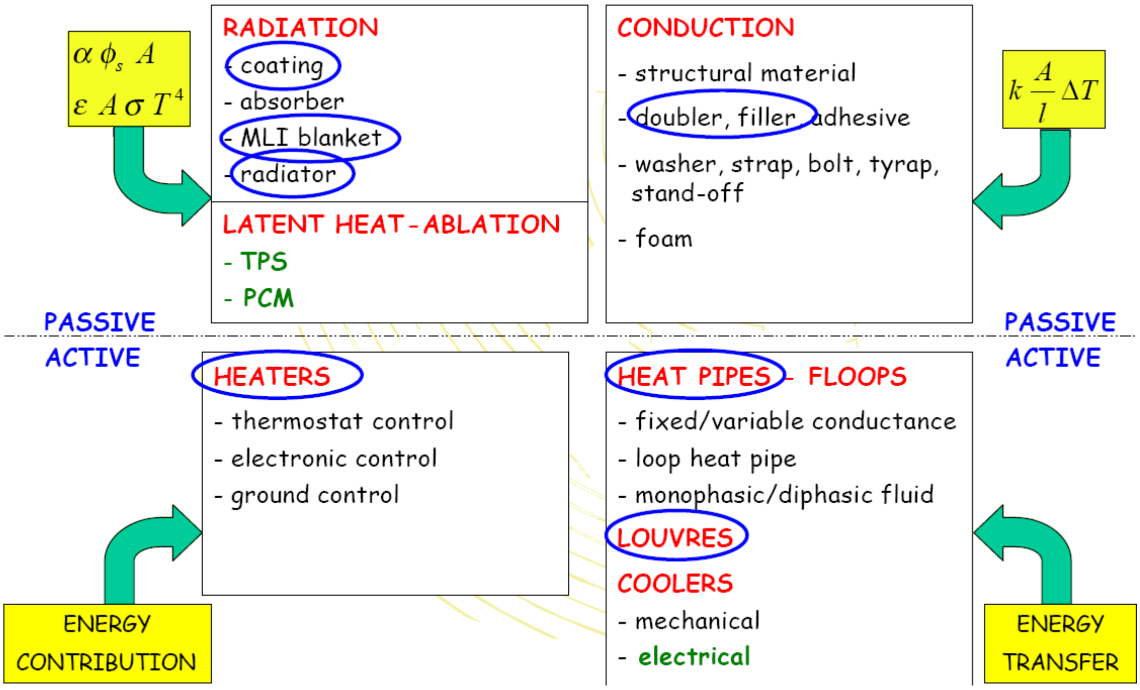
  – **heat pipe (Al tube)**

- 11 kW @20°C l=0.70 m Ø=2.5 cm DT= 3 K m= 2 kg

  – **radiation (from a black surface ε=1)**

- 11 kW @20°C A= 88 m^2 DT= 25 K
- 11 kW @20°C A= 27 m^2 DT= 290 K

Means



Means

Passive Systems Pros/Cons

– no mechanical moving parts or moving fluids, no power consumption

- simple to design/implement/test
- low mass and cost
- highly reliable

– BUT low heat transport capability

- except heat pipes

Active Systems Pros/Cons

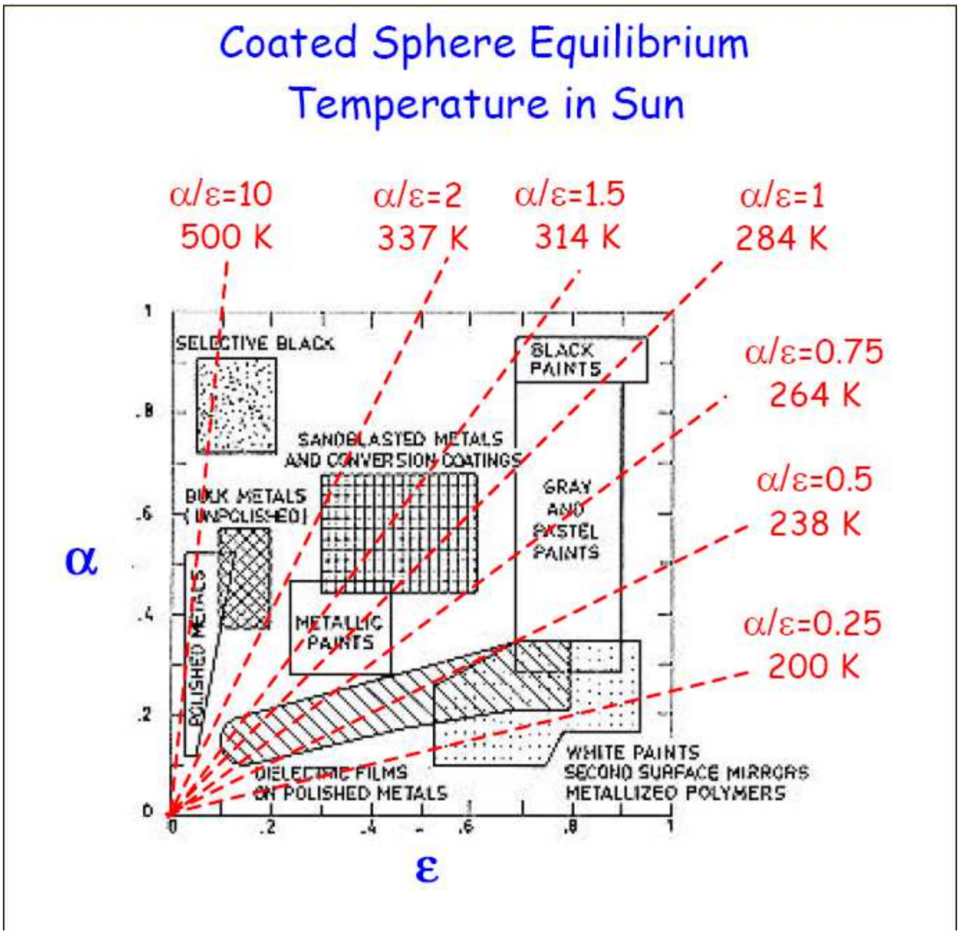
– mechanical moving parts or moving fluids or electrical power required

- complex design
- generate constraints on S/C design and test configurations
- high mass and cost

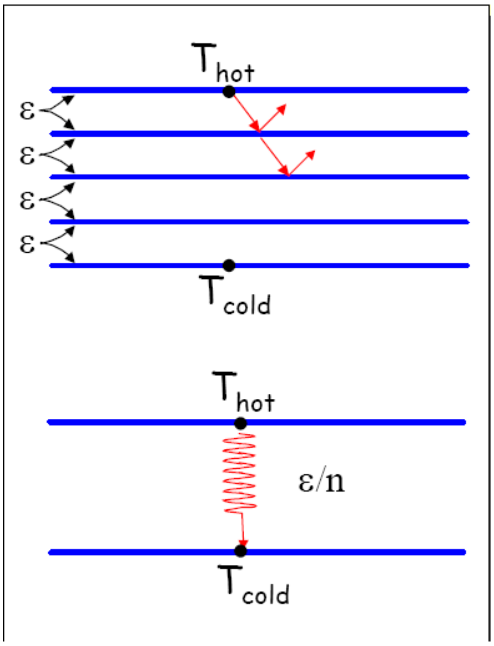
• less reliable than PTC means

Means – PTC – Radiation, coatings

- controls Heat absorbed by External S/C Surfaces
- with α, solar absorptivity
- controls Heat radiated to Space
- with ε, IR emissivity



Means – PTC – Radiation, MLI Blankets



- **Purpose**
- insulating material
- acts as a radiation barrier
- decreases heat flow inside S/C
  - Sun, albedo, IR planet
  - ascent aerothermal after fairing jettison
  - ME/ABM firing
- decreases heat losses from S/C
  - IR energy
- **Principle**
- stack of n layers with low emissivity e
- connected only by radiation with limited contact areas equivalent to reduce the emissivity by n.

Means – PTC – Radiation, Radiators

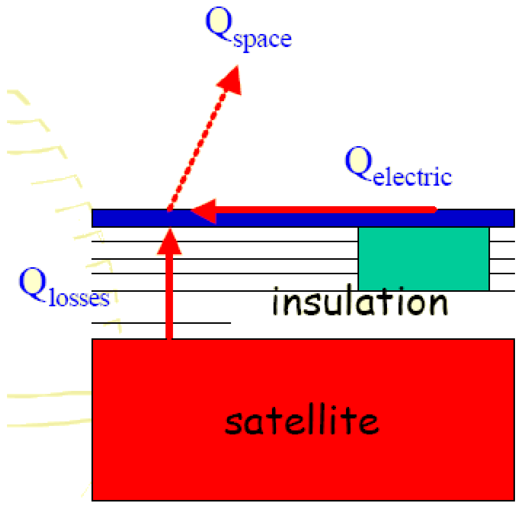
- **Purpose**

– cool detectors, optical components, mirrors

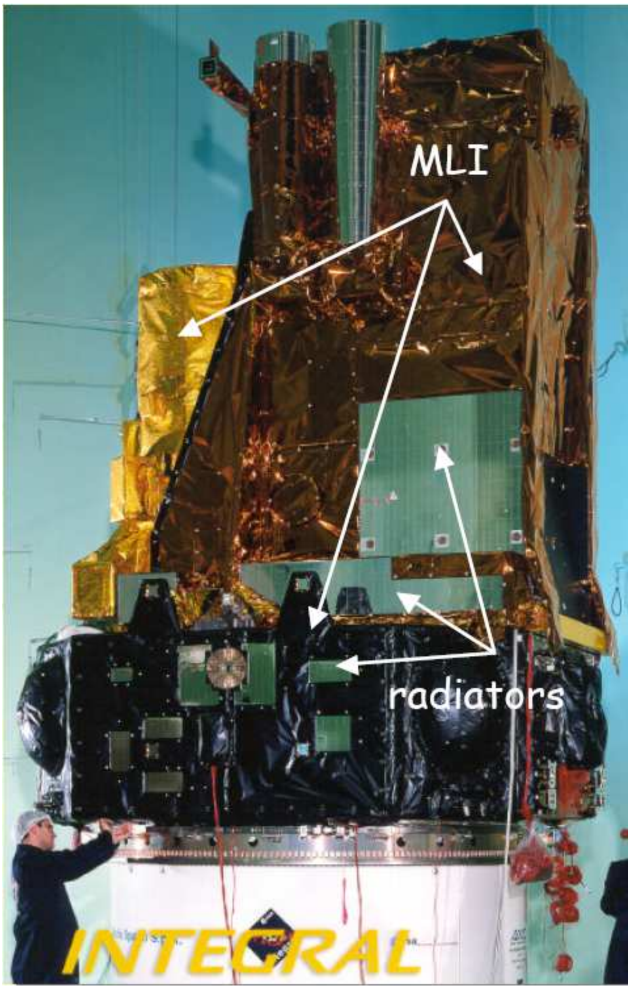
– improve the performances of

- Scientific P/L (all wavelength ranges)
- Earth observation (mainly IR)
- **Principle**
- direct coupling to deep space @2.73 K
- heat lift decreases in T^4 6 W/m2 @100K

$$Q_{space} = Q_{electric} + Q_{losses}$$



Q_{space}
$Q_{electric}$
Q_{losses}
insulation
satellite

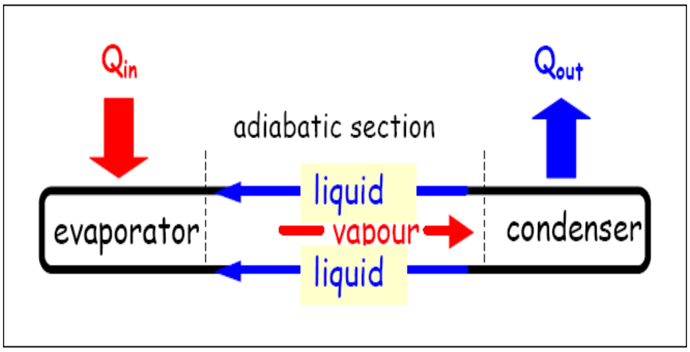


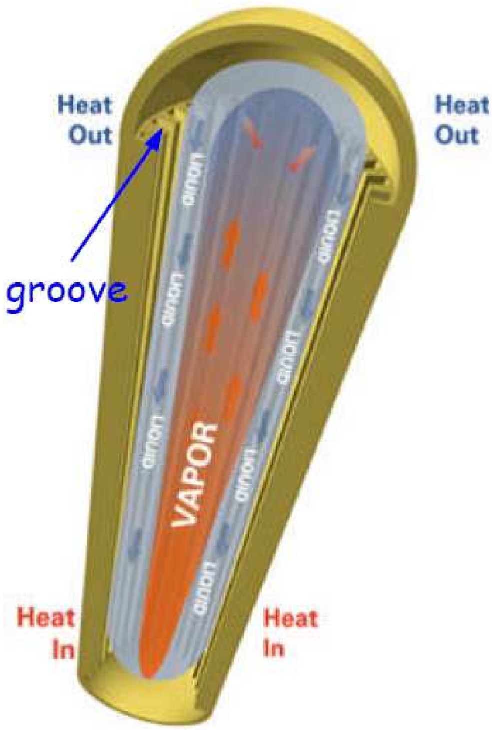
MLI
radiators
INTEGRAL

INTEGRAL STM

Means – PTC – Radiation, Radiators

- Purpose
- transport heat by convection with small DT
- avoid temperature gradients
- Principle
- liquid vaporizes at evaporator
- gas flows to cold end
- gas condensates at cold end
- liquid returns by capillary forces





Grooved Heat Pipe

Chapter 9: Launch Opportunities and Deployment

Launch and Deployment Chapter

Launch

In this section, the launch part will be covered. Admittedly, for Nano-satellites to perform the task they are required to do, they need to be launched in a launcher to space. It is worth mentioning that Nano-satellites are classified as satellites that are 1kg-10kg (wet mass= including fuel). These sections will discuss the following:

- Different types of launchers
- Different orbits
- Launch services and opportunities
- Launch costs

9.1 Different types of launchers

In this section, different types of launchers will be discussed that are used to launch Nano-satellites into space. There are several launchers which launched small satellites, including Nano-satellites into space. As mentioned earlier, Nano-satellites are classified as satellites that are 1kg-10kg (wet mass= including fuel). Figure 1 and Table 1 show some of the launchers than launch small satellites. The following includes the Nano-satellite ranges:

Vehicle/Configuration	Capacity to 700 km Sun-Synchronous Orbit (kg)	Cost ($M)	Performance History (Successes/Flights through Oct. 1988)
U.S. LAUNCH VEHICLES			
Delta II 7920/25	3,275	50	47/49
Delta II 7320	1,750	35	0/0
Pegasus XL	225	14	18[a]/23[b]
Taurus XL/Orion 38	945	24	0/0
Taurus/Orion 38	860	22	3/3
Athena 3	2,200	30	0/0
Athena 2	700	22	1/1
Athena 1	200	16	1/2
Conestoga 1229	220	12	0/0
Conestoga 1620	540	18	0/1
FOREIGN LAUNCH VEHICLES			
CZ-2D (China)	1,200	20	5/5
PSLV Mk2 (India)	1,300	12/15	1/1
Molniya M (Russia)	1,775	30	256/289
Shavit 2[c] (Israel/US)	340	15	0/0
Shtil 1N (Russia)	185	5/6	1/1
Tsyklon 3 (Ukraine)	2,300	25	111/117

[a] Successes exclude incorrect orbit, failure to separate on orbit, and damaged spacecraft.

[b] Includes all versions of the Pegasus.

[c] Coleman Research Corporation, in collaboration with Israel Aircraft Industries, has recently won a Small Expendable Launch Vehicle contract from the National Aeronautics and Space Administration to provide launch services in the United States using an export version of the Israeli-designed Shavit rocket (Next). The Shavit is a solid-fuel rocket with performance comparable to the Pegasus XL. Through January 1998, it had achieved three successful launches in five attempts.

SOURCE: International Space Industry Report, Nov. 9, 1998; available online at <CS:WebLink>http://www.launchspace.com/isir/home.html>

Fig 1. The National Academic Press (2017)

Table 1. Parabolic Arc (2016)

System Performance, Launch Type & Location				
Vehicle Name	Performance	Orbit	Launch Type	Launch Location
Devon Two	4 kg	LEO	Land	
Cab-3A	5 kg	400 km	Air	Int'l Water
Spyder	8 kg	370 km	Land	Wallops,Cape Canaveral
Bagaveev	10 kg	SSO	Land, Water	
Таймыр-1А	12 kg	LEO	Land	
North Star Launch Service	20 kg	350 km	Land	Andøya Rocket Range, Norway
NEPTUNE N5	40 kg	310 km SSO	Land, Water	Moody Space Centre, Australia. Int'l Water
Wolverine	45kg 35 kg	LEO SSO	Land, Sea	Kodiak,Cape Canaveral

Neutrino I	50 kg	LEO	Land	
Primo	50 kg	700 km SSO	Land	
Sagitarius Space Arrow	64 kg	600 km	Air	Int’l Water, Spanish airport
Bloostar	100kg 75 kg	LEO600 km SSO	Balloon	Int’l Water
Arion 2	150 kg	400 km	Land	South Europe
Electron	150 kg	500 km SSO	Land	Birdling’s Flat, New Zealand
Demi-Sprite	160 kg	LEO	Land	
LauncherOne	200 kg	SSO	Air	Int’l Water
Volant	215 kg	LEO	Land	Kodiak
Prometheus-1	250 kg	LEO	Unknown	
SOAR	250 kg	LEO	Air/Space	Canary Islands, Spaceport Colorado
Tronador II	250 kg	600 km SSO	Land	Puerto Belgrano Naval Base
Intrepid-1	376 kg	500 km SSO	Land	Kennedy Space Center
Firefly α	400kg 200 kg	LEO SSO	Land	Kodiak Preferred
Haas 2C	400 kg	LEO	Land	Spaceport America
Fei Tian 1	430 kg	500 km SSO	Land	
M-OV	454 kg	LEO	Land	
Black Arrow 2	500kg 200 kg	200km 600 km SSO	Land	Worldwide
Pegasus XL	468 kg	200 km	Air	Int’l Water; Multiple locations demonstrated
Orbital 500	500 kg	500 km SSO	Air	
LandSpace-1	530kg 400 kg	300km 500 km SSO	Land	China

The launchers mentioned in Figure 1 and Table 1 are just some of the launchers that are active now. This does not include the launchers that are in development and are expected to be in use in the short term. Table 2 shows some of the launchers which are in development and are expected to be in operation soon.

Table2. Parabolic Arc (2016)

Small Space Launch Systems Under Active Development				
Organization Name	Vehicle Name	Country	Latest Launch Date	Launch Frequency
PLD Space	Arion 2	Spain	2021	10/year
Bagaveev Corporation	Bagaveev	USA		50/year
Horizon Space Technologies	Black Arrow 2	United Kingdom		4/year
zero2infinity	Bloostar	Spain	2017	
CubeCab	Cab-3A	USA	2018	
Scorpius Space Launch Company	Demi-Sprite	USA		
Tranquility Aerospace	Devon Two	United Kingdom		
Rocket Lab	Electron	USA/New Zealand	Q4 2016	1/week
China Aerospace Science and Industry Corporation	Fei Tian 1	China		
Firefly	Firefly α	USA	2017	50/year
ARCA Space Corporation	Haas 2C	USA	2016	
Rocketcrafters	Intrepid-1	USA	Q4 2018	
LandSpace	LandSpace-1	China	Q4 2017	
Virgin Galactic	LauncherOne	USA	H2 2017	24/year
MISHAAL Aerospace	M-OV	USA		
Interorbital Systems	NEPTUNE N5	USA	2017	
Open Space Orbital	Neutrino I	Canada		
Nammo	North Star Launch Service	Norway	2020	
Orbital Access	Orbital 500	United Kingdom	2020	
Leap Space	Primo	Italy		20/year
SpaceLS	Prometheus-1	United Kingdom	Q4 2017	
Celestia Aerospace	Sagitarius Space Arrow	Spain	2016	
Swiss Space Systems	SOAR	Switzerland	2018	
UP Aerospace	Spyder	USA		

CONAE	Tronador II	Argentina	2019	
VALT Enterprises	VALT	USA		1000s/year
Bspace	Volant	USA	2018	Multiple/quarter
Vector Space Systems	Wolverine	USA	2018	100/year
Lin Industrial	Таймыр-1А	Russia	Q1 2020	

9.2 Different Orbits

Satellites are launched into different orbits for different reasons. This section will mention some of the most well-known orbits where satellites are launched.

The orbits where satellites are launched to orbit around are:

- Low Earth Orbit (LEO)
- Medium Earth Orbit (MEO)
- Geostationary Earth Orbit (GEO)
- Molniya/Highly Elliptical Orbit (HEO)

Figure 2 shows how the different orbits look around the Earth

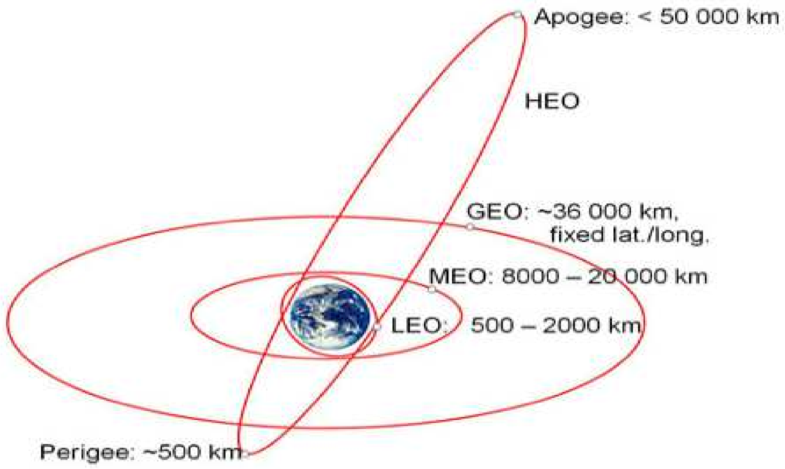


Fig 2. Computational Physics, Inc. (2014)

Each of these have their own characteristics, advantages and disadvantages. This is further detailed in Table 3 below.

Table 3. Characteristics of different orbits

Satellite feature	LEO	MEO	GEO	Molniya/HEO
Full Form	Low Earth Orbit Satellite	Medium Earth Orbit Satellite	Geostationary Earth Orbit Satellite	Highly Elliptical Orbit
Orbital Period	10min-2hrs	2hrs-24hrs	24hrs	24hrs
Satellite Height	150km-2000km	2000km-35786km	35786km	>35786km
Satellite Life	Short	Long	Long	Long
Propagation Loss	Least	High	Highest	High to medium
Advantages	•LEO satellite provides better signal strength. •Least signal propagation delay since it is closest to earth	•With MEO satellite lesser number of satellite network required compared to LEO satellite to cover the area. •Lesser time delay in the signal compared to GEO satellite	•Covers large geographical area, only three GEO satellites are needed to cover earth. •Visible for 24 hours from fixed location on earth •Ideal for satellite broadcast and multipoint communication	•Used for specific purpose by USSR to cover inaccessible area at higher latitude. •Molniya satellite is less costly to place in the orbit compared to GEO satellite •Molniya satellite is Very useful in the polar region
Disadvantages	•A very large number of satellites network is required; thus, it is very costly •Atmospheric	•Weaker signal compared to LEO since it is higher orbit than LEO	•Considerable time delay in the signal, which is not favorable for point to point communication. •Since it is above	•Ground station needs steerable antenna to track the spacecraft. •Satellite remains fixed for only 8 hours

	drag effects are more which cause gradual orbital disorientation •only visible for 15 to 20 minutes from an area	•Visible for only 2 to 8 hours from a location on earth	the equator, it faces difficulty in broadcasting near polar region	relative to the earth

9.3 Launch costs

The launch costs for Nano-satellites can vary depending on the mass of the Nano-satellite as well whether the satellites can be launched with another payload.

Table 4 gives an idea of the cost of launching small satellites. The costs depend on the mass and on the type of launcher.

Table 4. Parabolic Arc (2016)

Small Satellite Booster Launch Costs		
Vehicle Name	Projected Launch Cost	Estimated Cost per kg
Sagitarius Space Arrow	$0.2 M	$10.6 k
NEPTUNE N5	$0.5 M	$12.5 k
Black Arrow 2	$6.6 M	$13.2 k
Intrepid-1	$5.4 M	$14.4 k
Firefly α	$8.0 M	$20.0 k
Wolverine	$1.0 M	$22.2 k
Demi-Sprite	$3.6 M	$22.5 k
Unknown	$0.3 M	$25.0 k
Electron	$4.9 M	$32.7 k
Таймыр-1A	$0.5 M	$40.0 k
SOAR	$10.0 M	$40.0 k
Bloostar	$4.0 M	$40.0 k
Cab-3A	$0.3 M	$50.0 k
Launcher One	$10.0 M	$50.0 k
Spyder	$1.0 M	$125.3 k
VALT	$1.7 M	No mass spec

9.4 Launch services and opportunities

Unsually, Nano-satellites and small satellites are launched with other payload, often a big satellite. This is known as piggyback. The big satellite is then the main payload and the small or Nano-satellites are secondary or tertiary payload. Satellites can also go as a cargo. The launch costs do not include of the costs to build and test a Nano-satellites. Nano-satellites, like every satellite, go through different tests to ensure that they can withstand the launch and space environment. Some of the testing Nano-satellite go through are:

- Mechanical tests (to see whether they can withstand impact on the satellite)
- Vibration tests (to see whether they can withstand the launch)
- Environmental tests (to see whether they can withstand the extreme temperatures of space)

The launch services are provided by the company that owns the launcher or the company that has contracted the launcher to be used. For example, Ariane 5 is used by European Space Agency (ESA) but is contracted by Arianespace and built by Airbus Defense and Space.

Recently Indian Space Research Organization (ISRO) PSLV launcher launched 104 satellites at once, breaking a world record from the Russian space agency Roscosmos, which in 2014 launched 37 satellites at once. (Guardian, 2017)

Figure 3 shows a picture of the PSLV-C37 launch which took place in February 2017.

Fig 3. Spacenews (2017)

9.5 Nano-satellites strategy

The points below are useful for setting a strategy for your Nano-satellites. Disclaimer: The points below are not chronically ordered.

1. Define your mission statement/requirements

- What are your objectives
- What do you want to achieve?
- What instruments do you need?
- Mass/weight
- What height in LEO do you want to place your satellite

2. Define your budget

- Outline all the costs from design to launch and operation
- Create an inventory of all instruments and equipment needed

3. Define how many personnel you need to run the mission

- This will further help you in your budget

4. Have contingency plan

- What if the launcher and/or ground station you want is not available/accessible
- What if a delay happens in launch, communication with satellite once it is in space?
- How does that affect everything else?

5. Launch

- Is your satellite going as a piggyback or cargo?
- How does that affect previous points mentioned?

6. Deployment

- What deployment approach/technique would you use?
- How does that affect everything else?

7. What services (companies) are you planning to use?
8. Insurance, contracting and legal issues for your satellite

9.6 Deployment

This section will answer the following three questions:

- How are NanoSatellite deployed?
- Why are NanoSatellite deployed only in LEO?
- What are some of the companies/organizations that offer deployment services for nanosatellites and what are the methods that they use to achieve these services?

How are nanosatellites deployed?

Nanosatellites, unlike regular satellites, are typically 1 to 10 kg in mass (1). The size of nanosatellites, therefore, limits how they are deployed when compared to larger satellites. This may be easily demonstrated by a brief discussion on mechanisms used to deploy equipment on larger satellites (i.e., satellites greater than 500 kg in mass).

Regular satellites may be used for various mission types, and for most of these mission's solar arrays are deployed to allow the satellite to generate energy from the sun to power the internal components of the satellite.

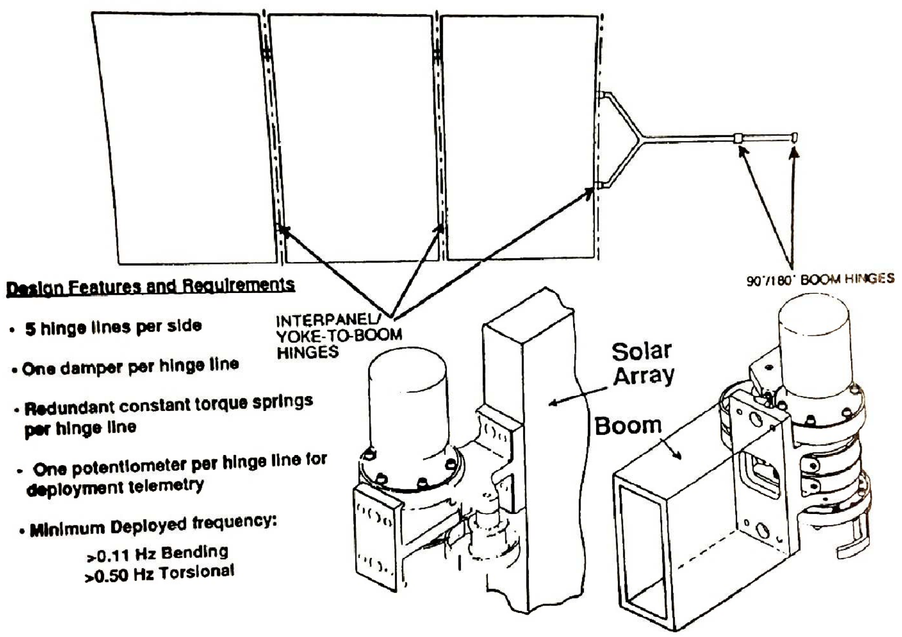


Figure 1: Typical Mechanism Locations on Satellite with Solar Array Configuration (2)

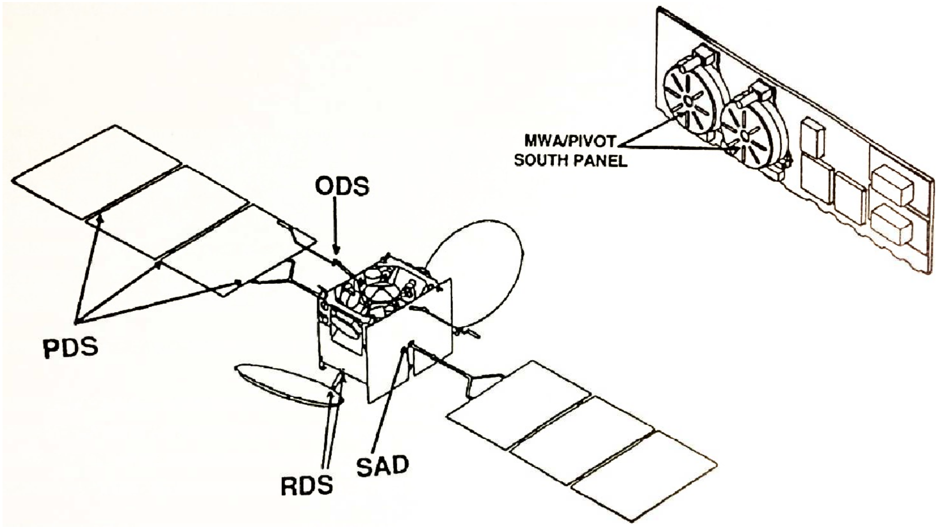


Figure 2: PDS Details (2)

Figure 1 shows a larger satellite, with its solar arrays deployed. A list of the acronyms in Figure 1 are described below.

MWA – Momentum Wheel Assembly, a continuously spinning wheel used to provide angular momentum for three axis spacecraft (2).

ODS – Omni Deployment System, an antenna reflector deployment hinge which is either spring or electrically driven.

PDS – Panel Deployment System, system that stows and releases solar arrays; see Figure 2 for a close of such a system.

SAD – Solar Array System, electrically driven system used on a three-axis spacecraft to rotate its solar arrays.

RDS – Reflector Deployment System, electrically driven hinge used to rotate communication antennas (see Figure 3).

Pyrotechnic cutters are sometimes used in a two-phase setup to deploy solar arrays; in Figure 3 pyrotechnic devices may be used to initially deploy the solar arrays and then the PDS used to complete the deployment of each solar array.

The use of nanosatellites has become popular, especially within the last decade (3). “This growth in the use of small satellites has been primarily driven by the miniaturization of electronics and sensors and the availability of commercial-off-the-shelf components with increasing capability, significantly reducing the cost of hardware development. The access-to-orbit and economy of these spacecraft is also improved through availability of secondary payload launch opportunities, especially for small satellites...” (3). However, such improvements have not resulted in nanosatellites typically capable of deploying large communication antennas and solar arrays (using the mechanisms) shown in Figures 1 through 3. Currently the Goddard Space Flight Center of the National Aeronautics and Space Administration (NASA) is in the process of designing a nanosatellite named Ice Cube, weighing 10 lbs., the size of a bread loaf, complete with a three-axis attitude control, deployable solar areas and a deployable communication antenna (4); see Figure 4.

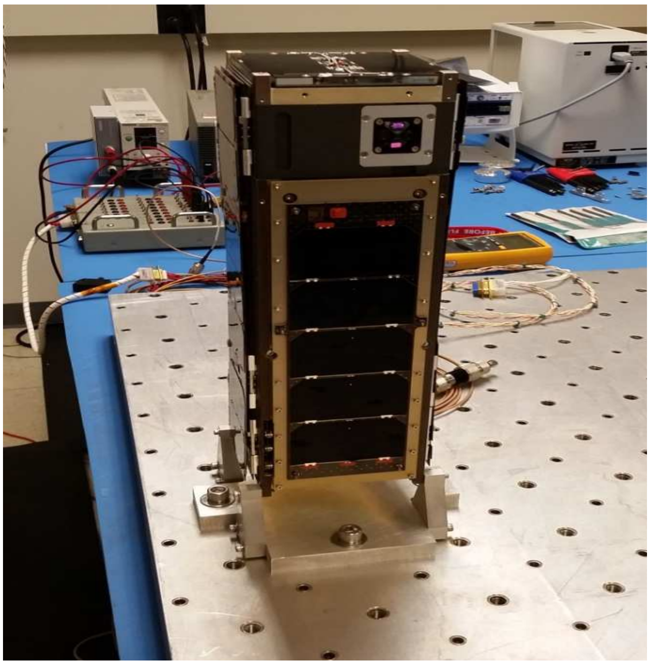


Figure 4: NASA’s IceCube Nanosatellite (4)

Due to the small mass and size of nanosatellites, the design of small thrusters with adequate thrust to control nanosatellites has been difficult and typically avoided due to safety concerns (5). Some examples of how nanosatellites are deployed, with the aforementioned considerations noted, are as follows:

- Multiple nanosatellites deployed at the same time, spread out over short distances to transfer data between each satellite. An example of this includes the two Nodes satellites deployed from the International Space Station (ISS) (6). These two satellites will eventually fall back down into the Earth's atmosphere and burn up.

- Individual nanosatellites deployed in orbit to measure the effect of microgravity on different materials and chemicals, aimed towards the Earth to take images to help meteorologists predict weather patterns or to study urbanization / deforestation to improve natural relief efforts and crop yields in developing nations (7).
- To receive navigation signals from aircraft. For example, the GomX-3, a follow-up to European Space Agency (ESA) 2013-launched Probe-V that first confirmed the feasibility of Automatic Dependent Surveillance – Broadcast (ADS-B) detection from orbit, was deployed in October of 2015 (8). This nanosatellite continues to help open the prospect of a global aircraft monitoring system that incorporates remote regions not covered by ground-based air traffic control (8).

9.6 Why are nanosatellites deployed only in LEO?

Nanosatellites are deployed only in LEO for some of the following reasons:

- Technology constraints; antennas and other components have not been miniaturized by space manufacturers yet to the point that they can carry out the same functions as those found on larger satellites. Nanosatellites are, therefore, limited in how far away they can be located from the Earth, to allow ground stations on our planet to transmit and receive data from them.
- Costs; the costs required to have a rocket to launch to an orbit greater than LEO, for any type of satellite, is significant. Accessibility for schools and universities to create and have nanosatellites deployed into space has prospered within the last several years, but only to the extent to allow such institutions to afford to have them placed into LEO.
- Nanosatellites are not designed with thrusters and any significant thrust applied would cause them to easily tumble out of control. Likewise, any other disturbances in space that can cause the nanosatellite to tumble out of control exist more the farther away it is from LEO. Nanosatellites in LEO are able to communicate to ground stations their status, even if in a tumbling state (if this is accounted for and/or important for the intended mission, of course). If deployed in

an orbit higher than LEO a nanosatellite could cause significant damage to other satellites if it went into a tumbling state.

What are some of the companies/organizations that offer deployment services for nanosatellites and what are the methods that they use to achieve these services?

"Together with NASA, companies like Orbital ATK, SpaceX, and NanoRacks give commercial companies the opportunity to fly their CubeSats as auxiliary payloads on cargo resupply missions to the International Space Station. In addition, Rocket Lab and Virgin Galactic will soon provide dedicated CubeSat launches from the new Venture Class Launch Services. CubeSats may be deployed directly from the rocket, from a spacecraft, or from the station itself depending on the mission..." (7).

Nanosatellites called CubeSats, are in the shape of cubes and tend to be easy to design and deploy. CubeSats have been a very popular type of nanosatellite and therefore are commonly used by both commercial companies and public institutions such as elementary schools and universities. Nanosatellites may be piggybacked on existing rockets carrying larger payloads (e.g., satellites) by companies such as Orbital ATK or SpaceX and then deployed into LEO before or after the main payloads have been deployed. This option tends to not be used often, due to the significant vibration loads that, like the main payload that the nanosatellite piggybacking must also carry. Nanosatellites are commonly sent to the ISS during cargo missions and then deployed by either being thrown into their planned orbit by the Japanese Experiment Module Robotic Manipulator System (a robotic arm, JEMRHS for short) (6), or if a CubeSat, launched out of the CubeSat Deployed on the ISS designed by NanoRacks LLC (9).

9.7 Launch and Deployment Strategy for Your Nanosatellite

The following points, in no particular order, should be considered when creating a new nanosatellite program:

1. Define your mission statement/requirements. What are your objectives? What do you want to achieve? What instruments do you need? Maximum desired mass? What height in LEO do you want to place your nanosatellite?

2. Define your budget (all mission components from design to launch and operation).

3. Define how many personnel you need to run all phases of the nanosatellite's life, ranging from its initial design all the way up to its mission.
4. Have a contingency plan. Consider whether you have access to a launcher and ground station that you need/want, if a delay happens and how it will affect everything else.
5. Launch (i.e., are you going to piggyback another satellite or are you going to use a cargo mission to the ISS?).
6. Deployment (what deployment approach/technique will you use).
7. What services (companies) are you planning to use.

References

[1] The National Academic Press, (2017), Small Launch Vehicles, https://www.nap.edu/read/9819/chapter/7

[2] Parabolic Arc, (2016), A Plethora of Small Satellite Launchers http://www.parabolicarc.com/2016/10/03/plethora-small-sat-launchers/

[3] Computational Physics Inc., (2014), Space Situational Awareness, http://www.cpi.com/capabilities/ssa.html

[4] RF Wireless World, (2012), GEO vs MEO vs LEO vs Molniya satellite, http://www.rfwireless-world.com/Terminology/GEO-satellite-vs-MEO-vs-LEO-vs-Molniya-satellite.html

[5] Spacenews, (2017), Countdown underway for launch for more than 100 satellites on an Indian Polar Satellite Launch Vehicle, http://spacenews.com/countdown-underway-for-launch-of-more-than-100-satellites-on-an-indian-polar-satellite-launch-vehicle/

[6] The Guardian, (2017), India launches record-breaking 104 satellites from single rocket, https://www.theguardian.com/science/2017/feb/15/india-launches-record-breaking-104-satellites-from-single-rocket

[7] Iowa State University, (2000-2001), Polaris Project Evening Star: Types of Orbit, http://www.polaris.iastate.edu/EveningStar/Unit4/unit4_sub3.htm

[8] NASA, (NA), Three Classes of Orbits, http://earthobservatory.nasa.gov/Features/OrbitsCatalog/page2.php

[9] Wikipedia.org, (2011). Small satellite. [online] Available at: https://en.wikipedia.org/wiki/Small_satellite [Accessed 5 Mar. 2017].

[10] Staley, D. (2003). AERO 4801 SPACECRAFT DESIGN. Carleton University Department of Mechanical & Aerospace Engineering

[11] Sciencedirect.com, (2015). Launch and deployment of distributed small satellite systems. [online] Available at: http://www.sciencedirect.com/science/article/pii/S009457651500171X [Accessed 5 Mar. 2017].

NASA.gov, (2017). Goddard's Ice Cube SmallSat Ready for Launch, Space Station [13] Deployment. [online] Available at:

https://www.nasa.gov/press-release/nasa-highlights-science-on-next-commercial-resupply-mission-to-international-space-0 [Accessed 6 Mar. 2017].

[14] Motherboard.vice.com, (2016). The Race to Build Teeny Tiny Satellite Thrusters. [online] Available at: https://motherboard.vice.com/en_us/article/the-race-to-build-teeny-tiny-satellite-thrusters [Accessed 6 Mar. 2017].

[15] NASA.gov, (2016). NASA Small Satellite Duo Deploys from Space Station into Earth Orbit. [online] Available at: https://www.nasa.gov/feature/nasa-small-satellite-duo-deploys-from-space-station-into-earth-orbit [Accessed 6 Mar. 2017].

[16] NASA.gov, (2017). CubeSats: Shaping Possibilities in Space. [online] Available at: https://www.nasa.gov/mission_pages/station/research/news/cubesats_possibilities [Accessed 6 Mar. 2017].

[17] Esa.int, (2015). ESA'S FIRST TECHNOLOGY NANOSATELLITE REPORTING FOR DUTY. [online] Available at: http://www.esa.int/Our_Activities/Space_Engineering_Technology/ESA_s_first_technology_nanosatellite_reporting_for_duty [Accessed 6 Mar. 2017].

[18] Wikipedia.org, (2011). NanoRacks. [online] Available at: https://en.wikipedia.org/wiki/NanoRacks [Accessed 6 Mar. 2017].

Chapter 10: Regulations and Policy

There is always regulations and policies that are necessary and mandatory to follow when designing and launching a NanoSatellite into space. Depending on where your satellite is developed and also where your satellite is going to be launched the policies and regulations must be met. Each country has different regulations you have to follow so it is very important to mission success that you keep in the regulations and do not break these regulations during your development and launch procedures. There are no universal and set regulations that you have to follow it all depends on what country, so you should research and know each countries regulations and policy. Since there are so many different regulations there is no way to cover all of them into one chapter in the NanoSatellite Engineering Professional certification, instead there is a list of regulation agencies and topics that need to be covered so you will know where to look and what to follow for your satellites mission because depending on your mission the regulations and policies will be different.

Since there are so many different regulations and policies from different countries below is a list of different resources you can view to make sure you follow these depending on where your launch and satellite development location is.

America

https://swfound.org/media/188605/small_satellite_program_guide_-_chapter_5_-_legal_and_regulatory_considerations_by_chris_johnson.pdf

http://www.space.commerce.gov/regulations/satellite-export-control-regulations/

http://www.itso.int/images/stories/Capacity-Building/Tunis-2015/Policy-and-Regulatory-Guidelines-for-Satellite-Services.pdf

http://www.aerospace.org/crosslinkmag/fall-2015/space-debris-mitigation-policy/

https://www.faa.gov/about/office_org/headquarters_offices/ast/legislation_policies/

http://www.americanbar.org/groups/young_lawyers/publications/the_101_201_practice_series/space_law_101_an_introduction_to_space_law.html

Canada

http://www.cba.org/Sections/Air-and-Space-Law

http://www.thecanadianencyclopedia.ca/en/article/air-law-and-space-law/

Europe

http://www.esa.int/About_Us/ECSL_European_Centre_for_Space_Law/Space_policy_documents_and_useful_readings_on_regional_and_national_space_legislations

http://spacegeneration.org/projects/space-law.html

http://www.esa.int/About_Us/Welcome_to_ESA/European_Space_Policy

Mexico

http://www.educacionespacial.aem.gob.mx/images/normateca_regulacion_cubesat.html

INDEX