Institute for
Communications
and Navigation

Satellite Constellations for Data Transfer from the Moon

Bachelor Thesis

Institute for Communications and Navigation
Prof. Dr. sc. nat. Christoph Günther

Supervised by:
Prof. Dr. sc. nat. Christoph Günther

by:
Quirin Funke

Munich, December 22, 2009

Contents

1 Abstract

This thesis is a study of orbit constellations for a satellite communication network at the moon. The focus of the network is to transfer data between Moon and Earth. This is why the constellation does not need to provide constant full coverage. A periodic communication possibility is sufficient. A good orbit constellation shall minimise the maximum delay a user has to wait for transmission and need as few satellites as possible at the same time.

Besides the inspection of orbit constellations, a short essay about the communication protocols TCP/IP and DTN which can be used for such a satellite network is part of this thesis as well.

2 Deutsche Zusammenfassung

Diese Bachelorarbeit mit dem Titel ”Satellitenkonstellationen für den Datentransfer vom Mond” beschäftigt sich damit, Satellitenkonstellationen auf ihre Tauglichkeit für ein Satellitenkommunikationsnetz am Mond zu überprüfen. Ziel dieses Netzes ist der Datentransfer vom Mond. Daher wird keine konstante Abdeckung der gesamten Mondoberfläche vorausgesetzt. Es ist ausreichend, wenn Kommunikationsmöglichkeiten in regelmäßigen Abständen auftreten.

Bei der Untersuchung möglicher Satellitenkonstellationen wurde darauf geachtet, die Anzahl der benötigten Satelliten gering und gleichzeitig die Wartezeit, bis eine gewünschte Übertragung stattfinden kann, so kurz wie möglich zu halten. Dabei stellte sich heraus, dass eine Konstellation mit vier inklinierten und gegenläufigen Satelliten auf einer Höhe von 550km wie sie in Kapitel 4.2.3 unter *”Variant 3”* aufgeführt ist, von den untersuchten Konstellationen die besten Ergebnisse liefert.

Des Weiteren wurden eine Konstellation für ständige Komplettabdeckung sowie eine Möglichkeit, die Mondrückseite konstant zu versorgen, untersucht. Dies ist über einen Satelliten an dem Librationspunkt L2, der sich im Erde-Mond-Systems hinter dem Mond befindet, möglich. An diesem Punkt heben sich die beiden Gravitationskräfte von Erde und Mond mit der Fliehkraft auf. Die Untersuchungen an einem Orbit um diesen Punkt wurden mit Hilfe der Satellitensimulationssoftware STK angestellt.

Im letzten Kapitel wird das Gebiet der Kommunikationsprotokolle für solch ein Satellitennetz kurz angeschnitten. Mögliche Protokolle (TCP/IP und DTN) werden kurz erläutert und auf Routingstrategien kurz eingegangen.

3 Introduction

The year 1969 is often mentioned as the birthday of the internet. From 1969's point of view, its features such as decentralisation, inherent redundancy and self-organisation were most innovative. Thanks to these properties, even 40 years later the internet is still growing and has changed our everyday life completely, for example through online shopping, emailing and voice-over-IP services.

1969 was also the year when mankind made its first step on another celestial body in the Apollo 11 mission. In the same 40 years since then, the Moon became increasingly interesting to researchers and continues to be visited by multiple space probes, preparing for future manned missions to our celestial neighbour. It is obvious that communication with these spacecrafts is a mission-critical task as a failure will result in a total loss of scientific outcome. At the same time, however, there is the need to reduce the costs for such missions considerably. For these reasons, engineers are looking for ways to make communication connections between the Earth and Moon-bound spacecrafts not only even more reliable, but also cheaper at the same time. A solution for this desire can be found by introducing advanced techniques such as decentralisation, inherent redundancy and self-organisation - properties we do know from Earth's internet.

A network like this consists of multiple satellites, each of them acting as network node. There are communication connections not only between the probes and ground, but also between the satellites themselves. This shifts the mission-critical importance of dedicated connections to a far more flexible architecture, that is even able to reorganise itself in case of failure of some connections. Data packets can be routed to their destination across multiple satellites, avoiding broken links and keeping all participating satellites on line in the best possible way. Using this principle, it is even possible to maintain contact to a satellite that is behind the Moon, a rover surveying the far side of the Moon or even an astronaut with a low-power radio unit.

To prove the properties and the general feasibility of a lunar internet, the European Student Moon Orbiter, designed, built and operated by European students and planned to be launched in 2014, contains a communication experiment called the LunaNet. LunaNet's purpose is to test required technology for a space internet. In addition to the features well known from Earth's everyday internet, there are additional requirements originating from high bit error rates and high signal propagation delay typical for long-distance space communication links. For a grown out satellite communication system there is also the need for optimised satellite orbits around the Moon in order to ensure optimal coverage of the Moon and its surrounding space with communication services.

This bachelor thesis aims at the investigation of the properties of such networks, particularly in terms of optimal satellite constellations and the protocols suited to satisfy the increased needs of a lunar internet. In the first chapter several satellite constellations are examined and their use for a communication network evaluated and the second chapter is about the protocols. The problems and solutions that occur if the internet protocols are used in space and a new approach that shall solve these problems by default.

4 Orbits

The proper orbit constellation is very important for the success of a satellite communication network and depends on the coverage requirements. In comparison to traditional satellite communication networks where constant coverage of a specific area is required, the scenario considered here has completely different needs. It is sufficient that every spot on the Moon sees at least one satellite once during an appropriate time window, but this must be guaranteed. As the exact definition of the time window is not important for the result, the decision was made to use one revolution period of the satellites. This makes calculations easier, but the size of the time window varies depending on the altitude.

The second requirement on the constellation is to minimise the number of needed satellites. Therefore several possible constellations were studied and rated depending on the needed satellites and the worst case delays the spots on the Moon experience. The numbers of needed satellites were calculated analytically while the worst case delays were computed numerically using Matlab® and visualised in a delay map.

4.1 Delay Map

The delay map is a map of the Moon surface showing the worst case delays in colour code for a specified set of satellites. The delay of a transmission is caused by the time the sender has to wait until the next satellite comes into sight plus the time it takes the satellite to reestablish a connection to Earth. This last time is in most cases zero as the link to the earth is already set up. Only when the satellite is behind the Moon the transmission data must be stored in the satellite memory and send to earth when it rises out of the shadow again. The worst case delay for one spot is then the highest delay that occurs if the sender always wants to transmit by the time one satellite just disappeared.

Figure 4.1 is an example for a delay map of a constellation with three polar satellites. The delay at the poles shown here in blue is very low in comparison to the equator (yellow to red). Two of the satellites pass by during one revolution at different times, thus the time between two contacts is at the poles half the time as at the equator. The higher delay on the far side of the Moon, located in the center of the map and marked in red, is a characteristic all delay maps have. The black lines show the satellite ground tracks and the arrows mark their starting positions and flight directions.

The Matlab® code behind the delay map takes as input arguments the Kepplarian elements of the circular satellite orbits. Based on this the times, when a spot has a link to a

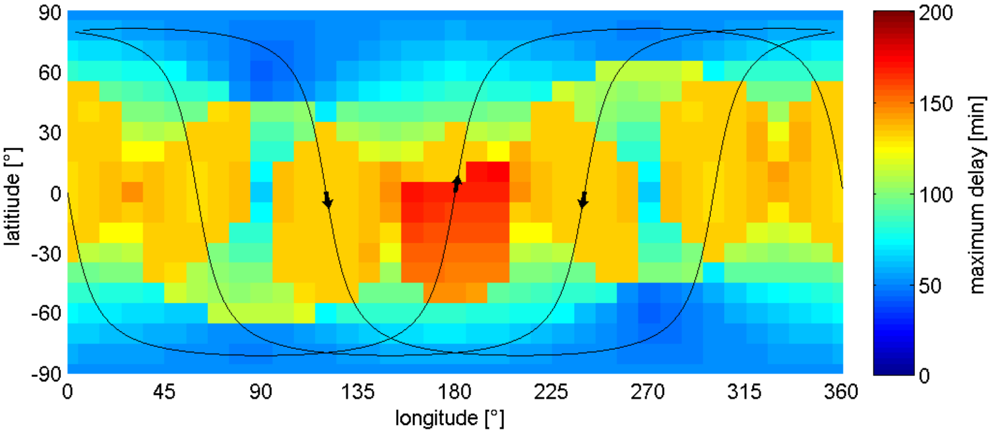


Figure 4.1: Delay map example. Blue areas mark low delays increasing over green and yellow up to red to the highest delays. The arrows mark the starting positions and flight directions of the satellites on their ground track visualised by the black lines. The altitude of the three satellites is 550km and their inclination 81°.

satellite are computed for each spot in the grid and each satellite in the constellation for a given simulation time. Also the starting times of links to earth are calculated for each satellite.

Now all these data are gone through for each spot and the delays for a transmission that starts right after one satellite disappeared are computed. The worst-case-delay that is shown in the delay-map then is the highest of these delays. The nearly equal additive link delay is not considered here.

As the libration of the Moon, a small movement of the earth aligned orientation due to eccentricity of it's orbit and the tilted rotation axis, is very small. It is not taken into account, because it would make the calculations more complicated without any real benefit.

4.2 Constellations

In order to find the best satellite constellation for this unusual scenario, common satellite constellations were studied and experiments with new ones done. The studied constellations are polar orbits, polar orbits with one equatorial orbit, inclined orbits and inclined orbits with one polar orbit. The last one did not show any benefit in the number of needed satellites and is therefore not mentioned here.

For analysis the number of needed satellites was computed analytically and the worst case delays examined using the delay-map. For the polar only and inclined orbits several variants of different flight directions and phases were studied and their delay maps computed. To make a fair comparison possible all delay maps were computed with four satellites at an altitude of 550km (see Chapter 4.2.4) and a minimum elevation angle for a successful contact of 5°.

4.2.1 Polar orbits

Needed satellites for full coverage

Polar orbits exclusively is probably the simplest constellation, but also the most inefficient as the coverage at the poles is very redundant.

To calculate the number of needed satellites we first need the half apex angle α of the coverage cone. It can be calculated using the law of sine:

$$\frac{\sin(\beta)}{R_M} = \frac{\sin(90^\circ + E)}{h + R_M} \tag{4.1}$$

where β is the half apex angle of the satellite beam cone, R_M the Moon radius, h the satellite height and E the smallest elevation where communication is possible. With

$$\alpha = 180^\circ - (90^\circ + E) - \beta \tag{4.2}$$

we get the following result for the half apex angle of the coverage cone:

$$\alpha = 90^\circ - E - \arcsin\left(\cos(E) \cdot \frac{R_M}{h + R_M}\right) \tag{4.3}$$

The criteria for full coverage is the coverage at the equator. Thus the number of needed satellites can be calculated as follows:

$$N_{pol} = \left\lceil \frac{180^\circ}{2\alpha} \right\rceil \tag{4.4}$$

The plot as a function of the height is shown in the combined Figure 4.18 in Chapter 4.2.4.

Constellation variants

Several variants of the constellation with quite different results in the delay map are possible. With the same number of satellites, orbital planes and altitudes, flight directions and phase can vary. This doesn't affect the coverage itself, but the maximum delay.

For polar orbits four variants have been studied (all in phase, with different flight directions, with different phases and with both together). It was tried to find the best constellation possible for all variants to make a fair comparison feasible.

Variant 1: All in phase

The variant with all satellites in phase (same flight direction and same phase) in Figure 4.2 is the simplest but also the worst in terms of delay as they all pass one line of latitude at the same time. So one spot that can communicate with multiple satellites sees them all at the same time but has to wait a full revolution period to the next contact.

As the satellites move northwards on the one side and southwards on the other side of the Moon, the areas around 157.5 and 337.5 longitude are passed twice per revolution which leads to a lower maximum delay at these spots (Figure 4.3).

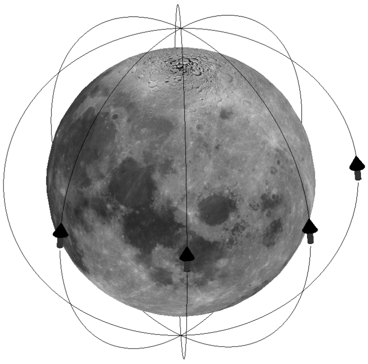


Figure 4.2: Constellation of four polar orbits with same flight direction and same phase at 550km altitude.

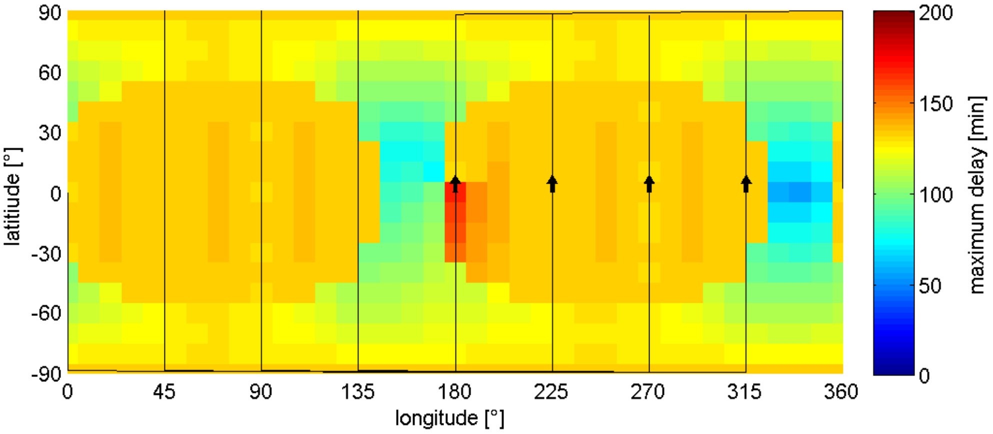


Figure 4.3: Delay map of four polar orbits with same flight direction and same phase at 550km altitude. The meaning of the elements is the same as in Figure 4.1.

Variant 2: Different flight directions

A method to improve the all-in-phase constellation is to change the flight directions of two of the satellites (Figure 4.4). Now the poles are passed twice per revolution but the equator is still passed just once per revolution. The delay at the areas around 157.5° and 337.5° is still better than in other regions because of the same reason as in variant 1. The resulting delay map is shown in Figure 4.5.

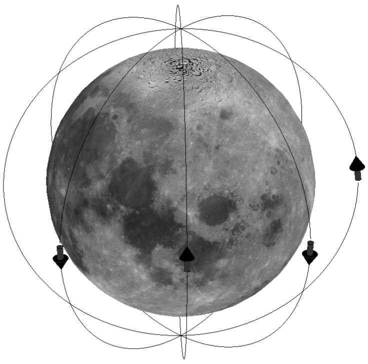


Figure 4.4: Constellation of four polar orbits with different flight directions and same phase at 550km altitude.

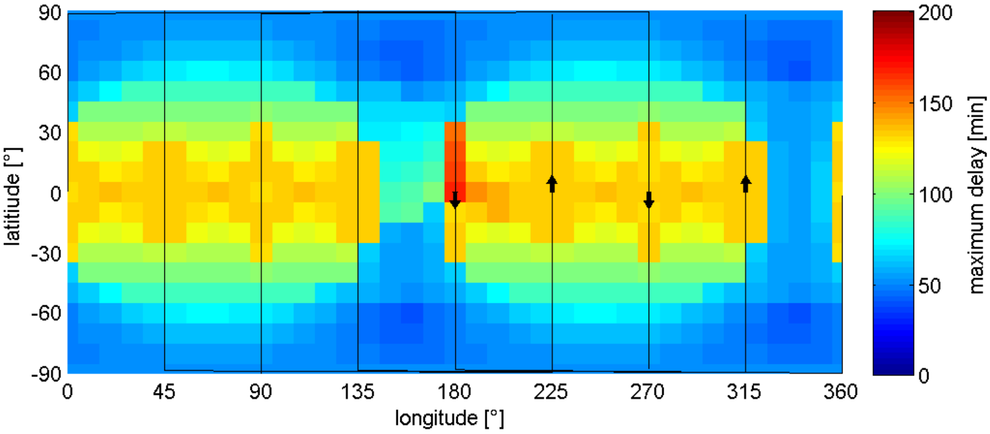


Figure 4.5: Delay map of four polar orbits with different flight directions and same phase at 550km altitude. The meaning of the elements is the same as in Figure 4.1.

Variant 3: Different phases

Another method to improve the all-in-phase constellation is to introduce a phase-offset between the satellites as shown in Figure 4.6. This solves the problem at the equator as adjacent satellites pass a spot which can see both of them, with a phase shift of half a revolution. That is way the spot has two contacts per revolution. The problem with four satellites is that this doesn't work out even, which leads to the two big spots with higher delay in Figure 4.7. An odd number of satellites would be better for this constellation.

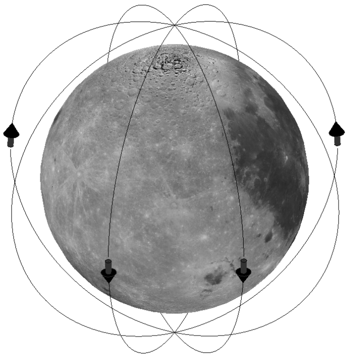


Figure 4.6: Constellation of four polar orbits with same flight direction and different phases at 550km altitude.

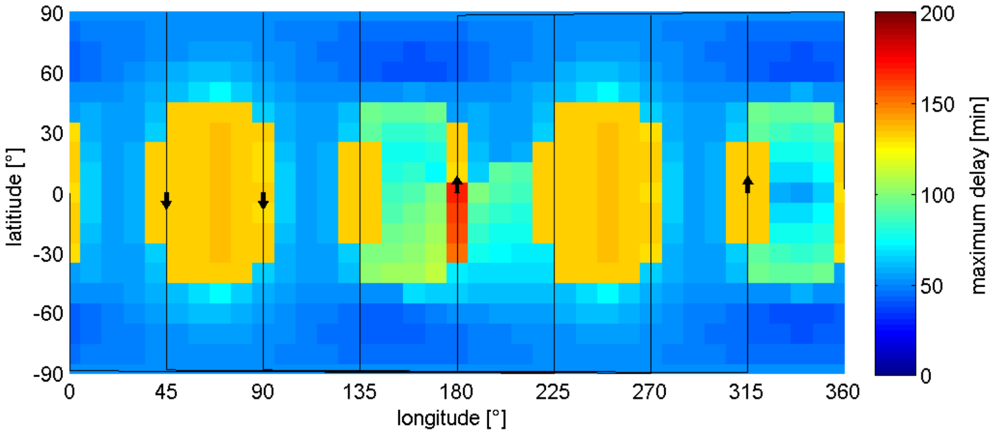


Figure 4.7: Delay map of four polar orbits with same flight direction and different phases at 550km altitude. The meaning of the elements is the same as in Figure 4.1.

Variant 4: Different phases and flight directions

A slight improvement of variant 3 is the combination of different flight directions and different phases. This constellation, shown in Figure 4.8, has a slightly lower delay at the high delay spots but also a slightly higher delay at some low delay parts of variant 3 (Figure 4.9). But the overall variance of the delay is smaller which leads to a more homogeneous coverage.

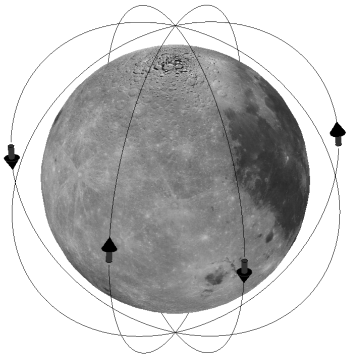


Figure 4.8: Constellation of four polar orbits with different flight directions and different phases at 550km altitude.

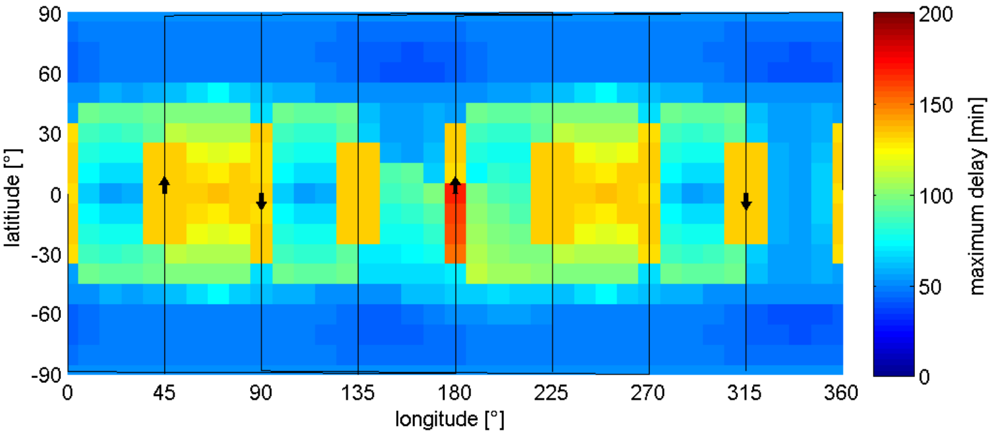


Figure 4.9: Delay map of four polar orbits with different flight directions and different phases at 550km altitude. The meaning of the elements is the same as in Figure 4.1.

4.2.2 Polar orbits with one equatorial orbit

Needed satellites for full coverage

The efficiency of polar orbits can be increased by adding an equatorial orbit to the constellation because the criterion for full coverage is not the coverage at the equator anymore but the coverage at latitude α (the half apex angle of the coverage cone of the equatorial satellite). This leads to a multiplication with the factor $\cos(\alpha)$ and one extra satellite:

$$N_{polequ} = \left\lceil \frac{180^\circ}{2\alpha} \cdot \cos(\alpha) \right\rceil + 1 \tag{4.5}$$

Constellation

Adding the equatorial orbit to the constellation has one huge benefit for the delay. The problem area around the equator gets additional coverage. For the polar satellites, only the best constellation from Chapter 4.2.1 was considered. This is the one with different flight directions and phases. As this constellation has an odd number of polar satellites the problem of the high delay spots does not occur anymore. The equatorial orbit is shifted about 180° in comparison to the polar orbits so that they complement ideally (Figure 4.10).

In the delay-map in Figure 4.11 it is shown that within the coverage band of the equatorial satellite the delay gets significantly better. Additionally the delay at the far side of the Moon is also much lower in this constellation.

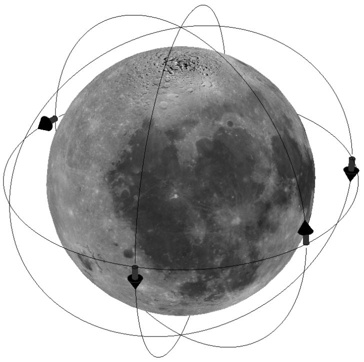


Figure 4.10: Constellation of three polar and one equatorial orbit with different flight directions and different phases at 550km altitude.

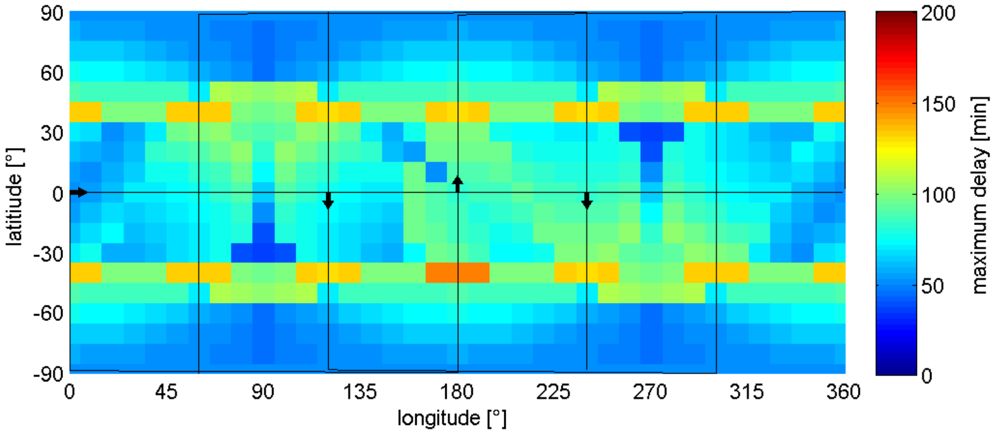


Figure 4.11: Delay map of three polar and one equatorial orbit with different flight directions and different phases at 550km altitude. The meaning of the elements is the same as in Figure 4.1.

4.2.3 Inclined orbits

Needed satellites for full coverage

Adding the equatorial orbit to the polar orbit constellation as done in Chapter 4.2.2 increases the efficiency but the coverage at the poles is still very redundant. The idea of the inclined orbits is to better disperse the coverage over the whole globe at the expense of the redundancy at the poles. To reach this goal the borders of the covered band of each satellite shall traverse at the poles. The criterion for full coverage is still the coverage at the equator, but the covered angle is bigger than 2α now, as the covered bands are rotated by α. Thus the new half of the covered angle at the equator is:

$$\sin(\alpha') = \frac{\sin(\alpha)}{\cos(\alpha)} = \tan(\alpha) \tag{4.6}$$

and the number of needed satellites can be calculated as follows:

$$N_{incl} = \left\lceil \frac{180°}{2\alpha'} \right\rceil = \left\lceil \frac{180°}{2 \cdot (\arcsin(\tan(\alpha)))} \right\rceil \tag{4.7}$$

Constellation variants

In this constellation are also several variants possible and have quite different results in the delay-map. The number of satellites, orbital planes and the altitudes stay the same again and flight directions and phases vary.

Just like for the polar orbits four variants were studied again: All in phase, with different phases, with different flight directions and both together. As the different flight directions lead to the best result and the additional different phases have no improvement, only the first three variants are mentioned.

It was tried to find the best constellation possible for each variant to make a fair comparison feasible here as well.

Variant 1: All in phase

The simplest variant is with all four satellites in phase again, meaning all having the same flight direction and phase as shown in Figure 4.12. The fourth satellite is behind the Moon because it wasn't possible to get all of them on one expressive picture.

With the inclined constellations the poles are the weak points and with this variant all satellites cover them at the same time. Thus the poles have only one contact to a relay satellite per revolution and because of this a high maximum delay. Figure 4.13 shows the belonging delay map. The situation at the equator is better, as two satellites pass by during one revolution. The far side of the Moon has a slightly higher delay because the satellites need to store the data until they get out of Moon's shadow again.

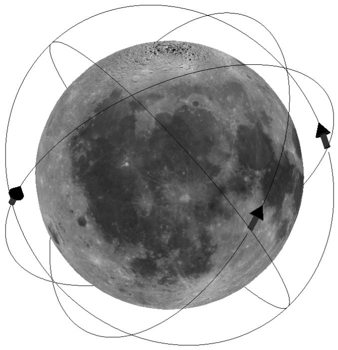


Figure 4.12: Constellation of four inclined orbits with same flight direction and same phase at 550km altitude.

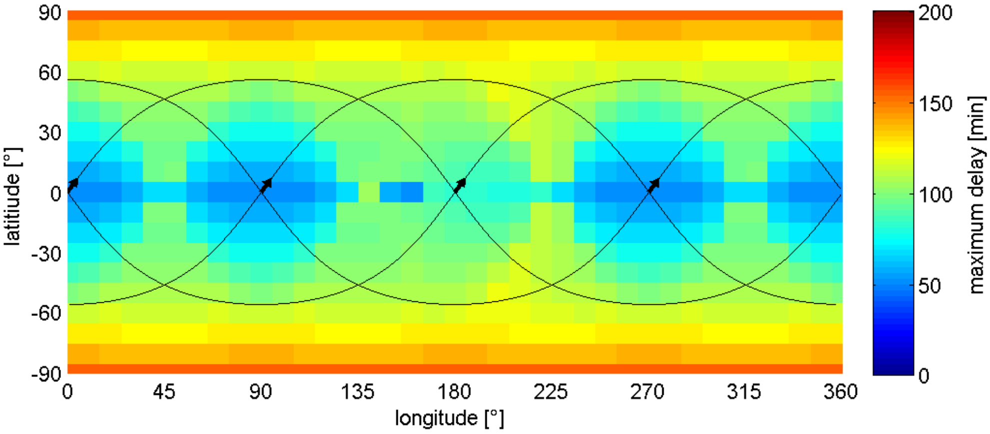


Figure 4.13: Delay map of four inclined orbits with same flight direction and same phase at 550km altitude. The meaning of the elements is the same as in Figure 4.1.

Variant 2: Different phases

To improve variant 1 a phase offset between the satellites can be introduced again which leads to the optimised constellation in Figure 4.14.

The delay map in Figure 4.15 shows a significant improvement at the poles which results from the different directions the satellites are crossing the equator now, two satellites northwards and two southwards. Thus the poles are covered by just two satellites at the same time now, leading to two contacts per revolution. The situation at the equator between two nodes is still a bit unhandy as two satellites are passing by, but with one fourth revolution difference only. This leads to a gap of a three fourth revolution afterwards.

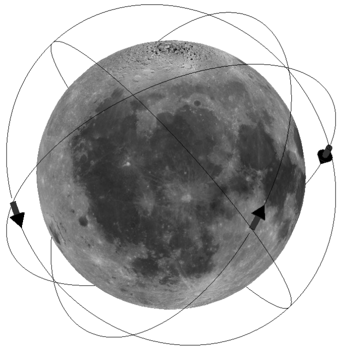


Figure 4.14: Constellation of four inclined orbits with same flight direction and different phases at 550km altitude.

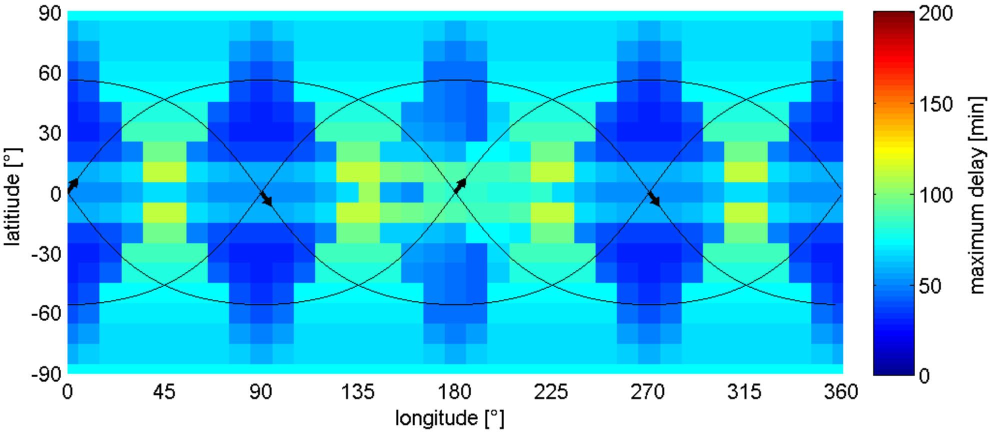


Figure 4.15: Delay map of four inclined orbits with same flight direction and different phases at 550km altitude. The meaning of the elements is the same as in Figure 4.1.

Variant 3: Different flight directions

The other method to improve variant 1 is putting the satellites on different flight directions as shown in Figure 4.16. As they cross the equator in different directions again, the problem at the poles is solved as well. In addition the problem spots of variant 2 at the equator between two nodes are fixed as the phase difference between two passing satellites is half a revolution. The delay map in Figure 4.17 shows this very well.

In this constellation the average delay and also the variance in delay are very low. Only the far side of the Moon experiences a slightly higher delay for obvious reasons. This problem will be solved with the Libration Point satellite in Chapter 4.3.2.

Experiments with a combination of different flight directions and different phases differences did not show any improvements. In fact these combinations mostly led to impairments. That is why it is not mentioned here anymore.

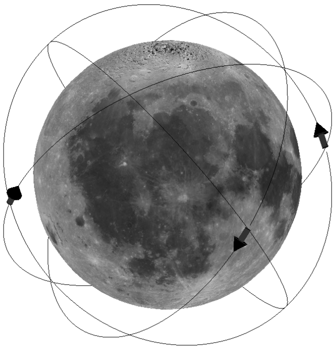


Figure 4.16: Constellation of four inclined orbits with different flight directions and same phase at 550km altitude.

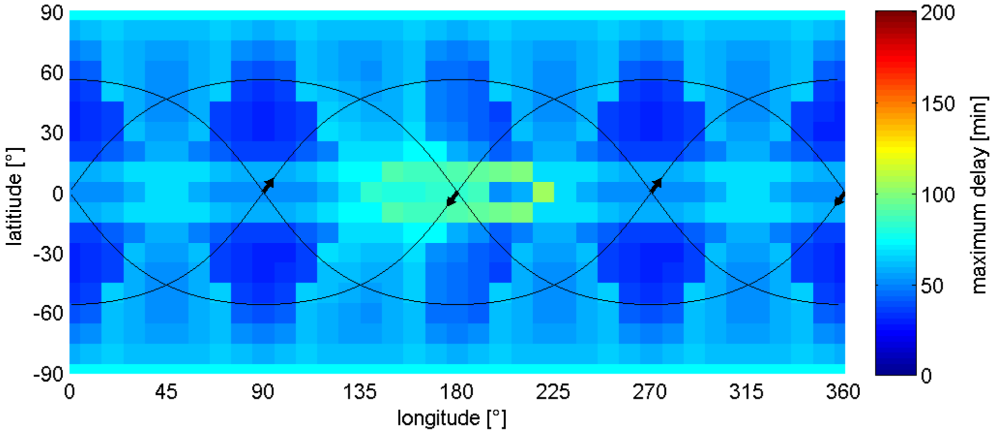


Figure 4.17: Delay map of four inclined orbits with different flight directions and same phase at 550km altitude. The meaning of the elements is the same as in Figure 4.1.

4.2.4 Comparison of the constellations

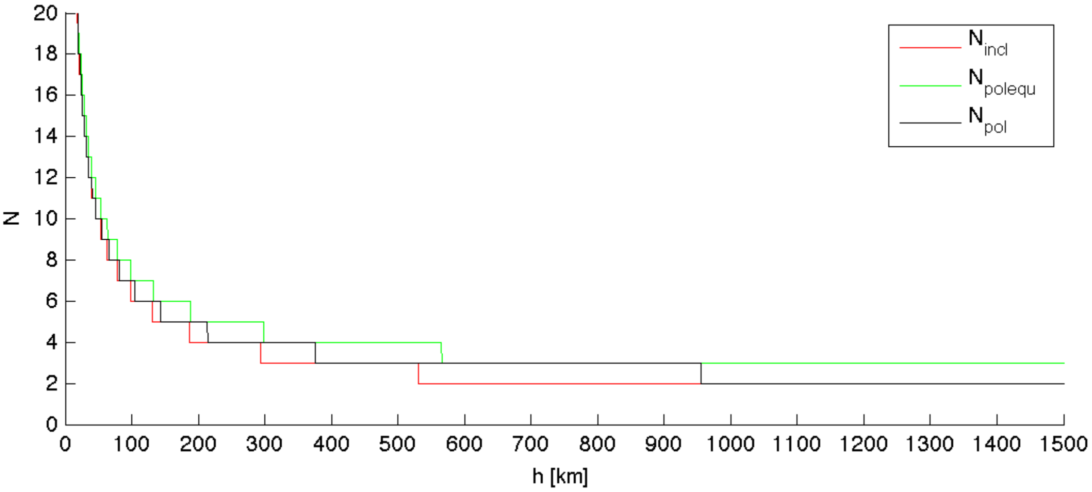


Figure 4.18: Comparison of the needed satellites of the three constellations. In black the polar constellations, green showing the polar with one equatorial constellation and red the inclined.

As Figure 4.18 shows, the constellation with inclined orbits is the most efficient. It always needs less or equal satellites as the other two constellations to cover the whole Moon surface. The polar with one equatorial satellite constellation is the worst in this discipline. It needs the most satellites to achieve full coverage.

But the number of needed satellites is not the only important attribute. The delay needs to be considered, too. The constellation with one equatorial satellite can score here in comparison to polar satellites only. But the inclined orbits are in this category the best again.

The computed delay maps are all without taking Moon's rotation into account and with a simulation time of one day. On this plot the regional coverage can be seen very good and bad covered areas marked in red can be located and eliminated by optimising the orbital parameters of the satellites. But to get a more realistic simulation Moon's rotation must be taken into account and the simulation time needs to be the revolution time of the Moon, 28 days. As this simulation takes quite a while it was only done with the best constellation, the inclined orbits with different flight directions.

At first the simulations were done with an altitude of 500km. When doing the simulation with Moon's rotation taken into account it revealed a weak point around the equator as Figure 4.19 shows it. Some spots come out of sight with the satellite due to Moon's rotation. Increasing the altitude to 550km solves this problem. The delay map in Figure 4.20 shows a nearly constant maximum delay with only the usual higher delay area at the centre.

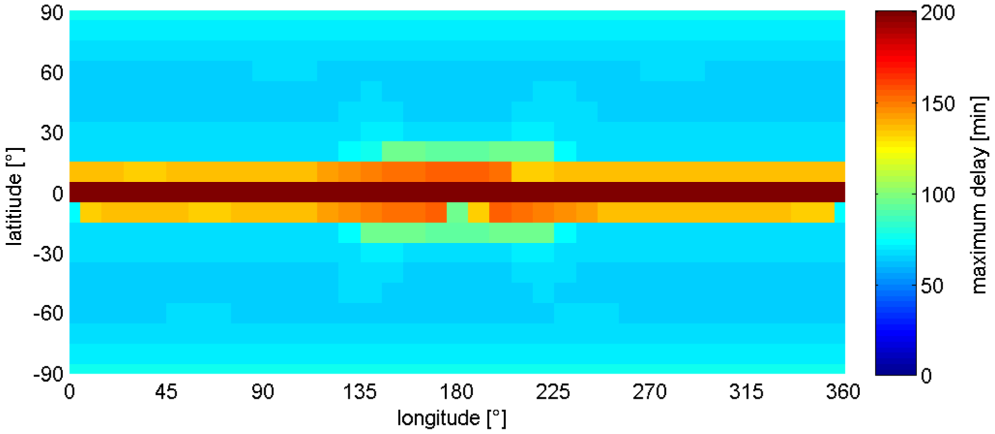


Figure 4.19: Delay map of four inclined orbits with different flight directions and same phase at 500km altitude with Moon's rotation taken into account and a simulation time of 28 days. The meaning of the elements is the same as in Figure 4.1.

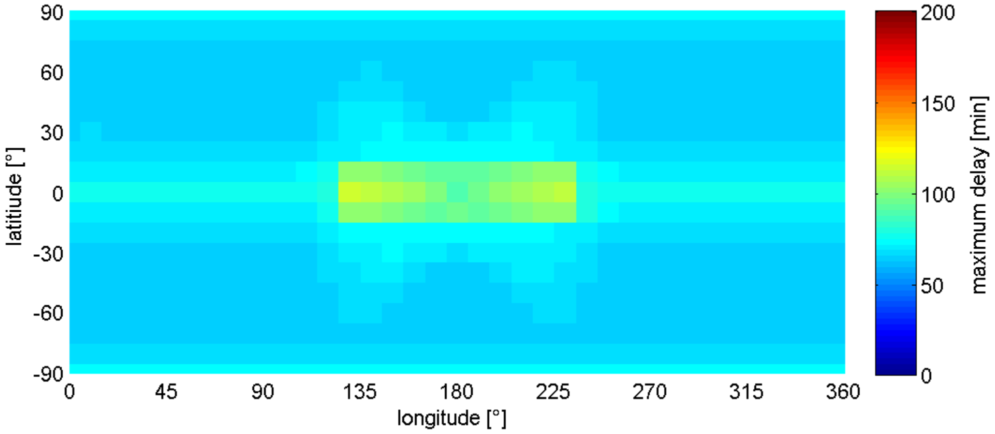


Figure 4.20: Delay map of four inclined orbits with different flight directions and same phase at 550km altitude with Moon's rotation taken into account and a simulation time of 28 days. The meaning of the elements is the same as in Figure 4.1.

4.3 Special constellations

This chapter is about two special constellations which cover areas of the Moon permanently, in contrary to the original LunaNet idea. A simple polar orbit constellation that covers the whole Moon surface and a constellation with only one satellite at the Libration point L2 that covers Moon's far side.

4.3.1 Constant coverage with polar orbits

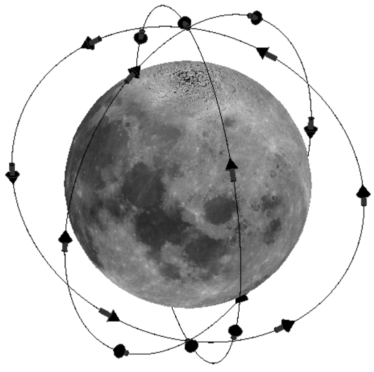


Figure 4.21: Example for a constant coverage constellation with three polar orbits planes and six satellites on each plane at an altitude of 800km.

In contrary to the previously considered constellations this one covers every spot on the Moon at all times. It is basically an extension of the constellation with polar orbits from Chapter 4.2.1. But now the equator needs to be covered by more than one satellite per orbit. Every orbit must be filled by enough satellites so that a constant coverage of the equator is assured. If the equator is covered constantly every spot on the Moon is covered constantly as the distance between two satellites on neighbouring planes is largest at the equator.

A small improvement can be made if the satellites on two neighbouring orbital planes are shifted in phase so that the overlap of the covered areas is minimised. The geometry of this shift is shown in Figure 4.22. Figure 4.21 shows an example constellation with three orbital planes and six satellites on each plane.

As the satellites' orbits all have the same characteristics, perturbations affect all satellites alike and the constellation keeps it's shape over time.

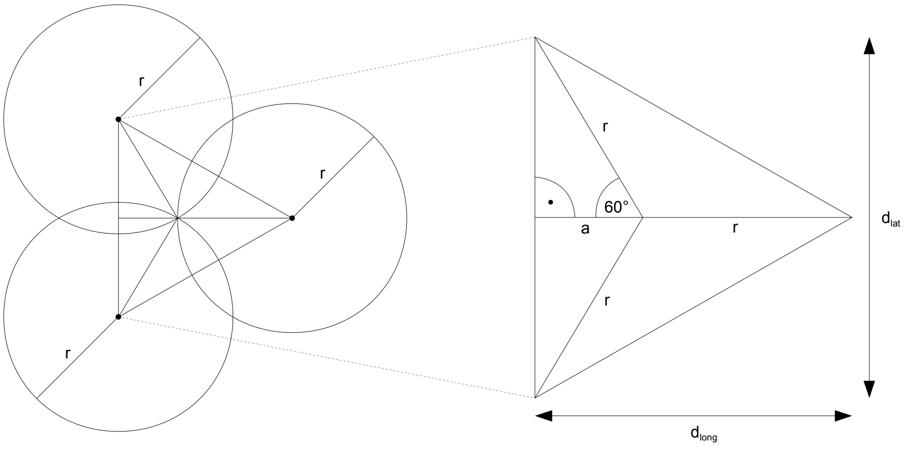


Figure 4.22: Geometry of the calculations on the constant link constellation. On the left side the covered areas of three neighbouring satellites are shown. On the right side a zoom into the interesting area for the calculation.

Needed satellites for constant coverage

The radius of the covered circle can be calculated out of the half apex angle α of the covered cone and Moon's radius R_M:

$$r = \alpha \cdot R_M \tag{4.8}$$

The distance d_{long} between two neighbouring orbit planes on the equator can be calculated out of r. As this distance is also the covered part of the equator of one orbital plane, the apex angle α_{long} of the covered part can be obtained using Equation 4.8:

$$d_{long} = r \cdot \cos(60°) + r = \frac{3}{2} \cdot r = \frac{3}{2} \cdot \alpha \cdot R_M \quad \Rightarrow \quad \alpha_{long} = \frac{3}{2} \cdot \alpha \tag{4.9}$$

The distance d_{lat} between two neighbouring satellites on one orbit plane can be calculated out of r, too. As this distance is also the covered part of one satellite on a line of longitude, the apex angle α_{lat} of the covered part can be obtained using Equation 4.8:

$$d_{lat} = 2 \cdot r \cdot \sin(60°) = \sqrt{3} \cdot r = \sqrt{3} \cdot \alpha \cdot R_M \quad \Rightarrow \quad \alpha_{lat} = \sqrt{3} \cdot \alpha \tag{4.10}$$

The orbital planes need to cover half of the equator as the other half is then covered automatically. The satellites on one orbital plane however need to cover the full circle.

The resulting equation for the number of needed satellites for a full constant coverage is:

$$N_{c,pol} = \left\lceil \frac{\pi}{\alpha_{long}} \right\rceil \cdot \left\lceil \frac{2\pi}{\alpha_{lat}} \right\rceil \tag{4.11}$$

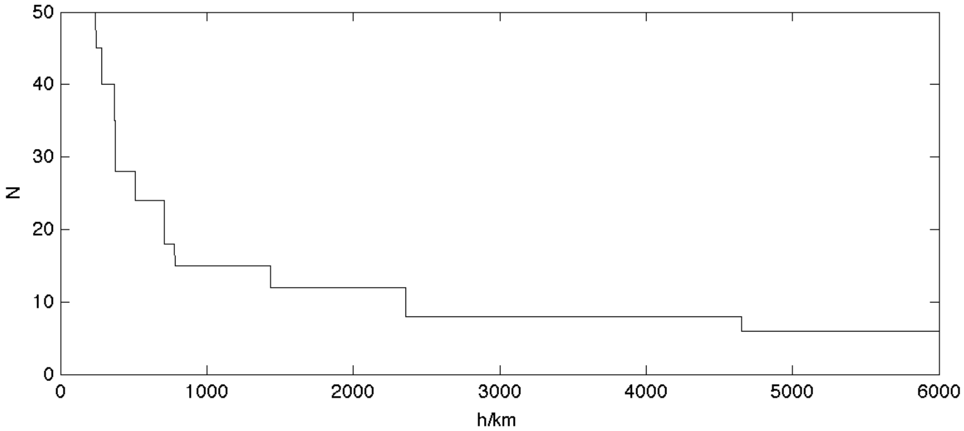


Figure 4.23: The number of needed satellites (N) for a constant coverage constellation with polar orbits as a function of the altitude.

Figure 4.23 shows the result from Equation 4.10 as a function of the altitude. The constellation with the fewest needed satellites is possible from an altitude of 4651km where two orbital planes with three satellites each are needed.

To make a constant connection possible even at the far side of the Moon, inter-satellite links are mandatory. Satellites in Moon's shadow need another satellite outside of it and in visibility to relay the data to Earth. But as the covered areas of neighbouring satellites do overlap a satellite can always reach it's neighbours.

4.3.2 Libration point orbit

In an orbital configuration with a central body and a smaller one orbiting it, like the Earth-Moon configuration, five points of orbital equilibrium, called Libration points, exist [7]. In these points the sum of the two gravitational pulls and the centripetal force equals zero.

L1, L2 and L3 are located on the line defined by Earth and Moon. L1 between them nearer to the Moon, L2 behind the Moon and L3 on Moon's orbit but on the opposite of Earth. They are all stable in the plane perpendicular to the Earth-Moon axis but unstable in the direction of the axis itself. It is possible to stay in an orbit around these Libration points but periodical correction manoeuvres are needed [7]. L4 and L5 are on Moon's orbit 60° ahead and behind the Moon. They are both instable in the orbital plane of Moon but the Coriolis effect keeps objects at these two Libration points on Halo orbits around them. Figure 4.24 shows the Libration points of the Earth-Moon system.

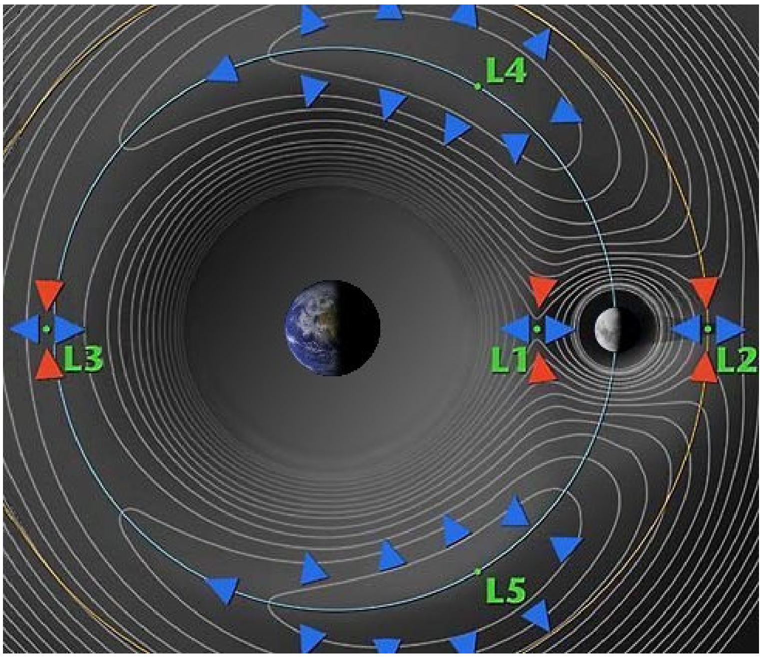


Figure 4.24: Position and stability of the five Libration points in the Earth-Moon system. Source: NASA

An orbit around the libration Point L2 has the positive attribute that the satellite is always in view of the far side of the Moon, which makes it very interesting for a communication satellite. The far side could constantly be covered and with a large enough orbit radius constant connection to earth is also possible.

As an analytical calculation of an orbit around the L2 is nearly impossible, the orbit in Figure 4.25 was computed numerically using *STK* of *AGI*. The shape of the orbit is not circular or periodic. The satellite is orbiting the L2 on a Lissajous curve which can be seen in Figure 4.26, showing the same orbit but for a time period of eight instead of two

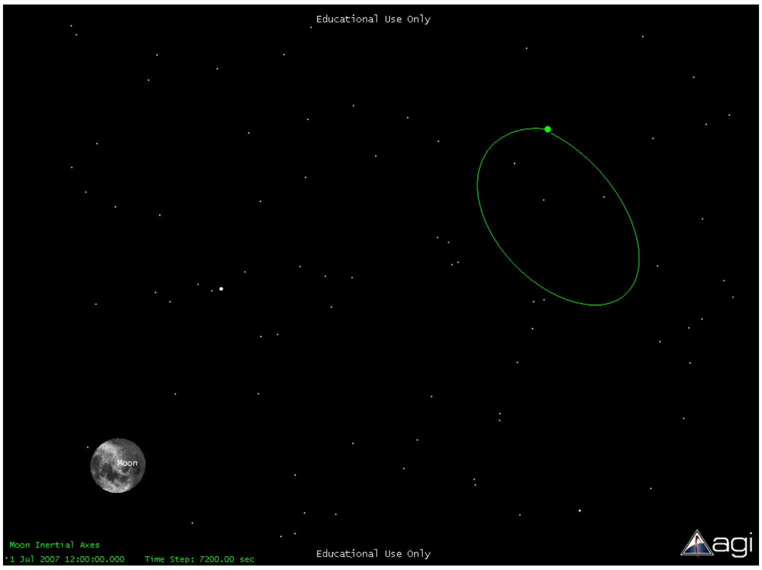


Figure 4.25: Orbit around the libration point L2 for a two week time period.

weeks. After about one year the Lissajous curve is repeating itself. Figure 4.27 shows the orbit for a one year time period. As the satellite gets very near to the center of it's orbit, the libration point, constant connection to earth is not guaranteed anymore. During the purple sections of the orbit, communication with earth is not possible (Figure 4.27).

In addition to the oscillations perpendicular to the Earth-Moon-axis responsible for the Lissajous curve there is a third oscillation in the direction of the Earth-Moon-axis forming the trajectory to a tube. Figure 4.28 shows detailed views of the constellation. The x-axis is defined as the Earth-Moon axis pointing away from Earth, the y-axis is the direction of motion of the coordinate system and the z-axis is perpendicular to the two others.

Figure 4.29 shows the proportions and distances of the constellation.

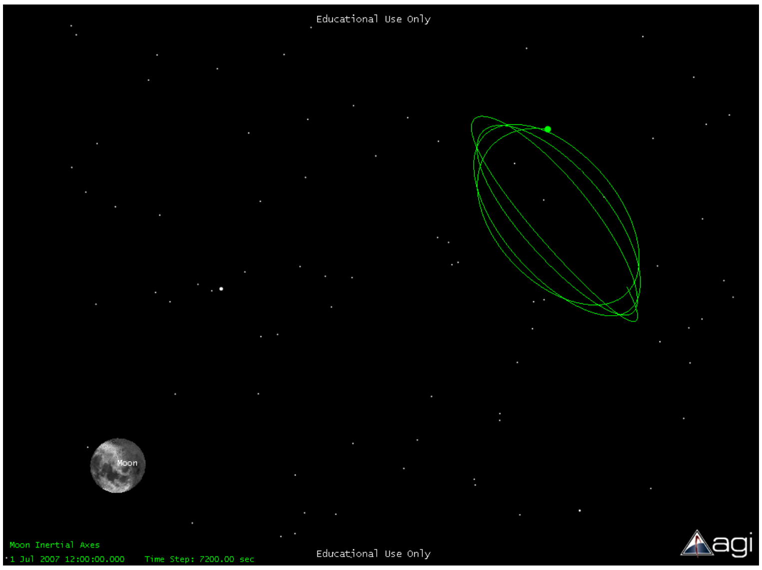


Figure 4.26: Orbit around the libration point L2 for an eight week time period.

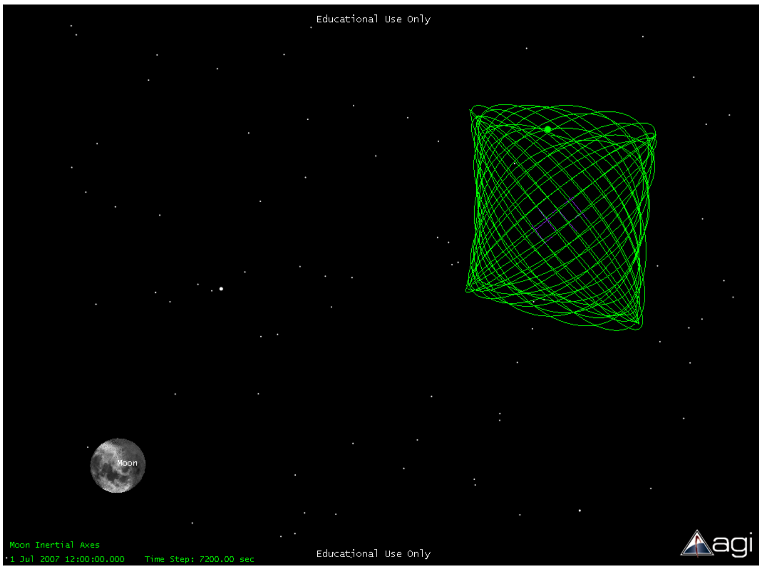


Figure 4.27: Orbit around the libration point L2 for a one year time period. The purple sections of the orbit mark the phases of no connection to earth.

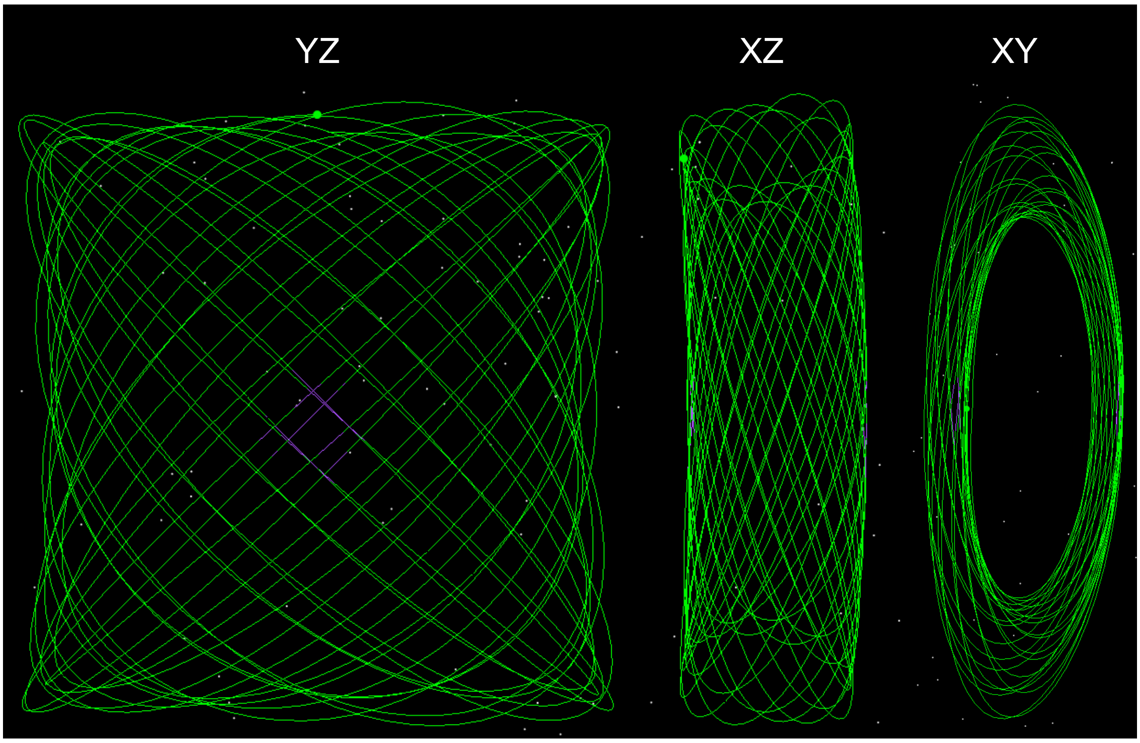


Figure 4.28: Detail views of the L2 orbit. Left: view of the YZ plane (as seen from Earth). Middle: View of the XZ plane (from the side). Right: View of the XY plane (from the top).

Figure 4.29: Wide view of the L2 constellation with Moon and Earth in the picture, illustrating the proportions.

Manoeuvres

As the Libration point L2 and orbits around it are instable in the direction of the Earth-Moon-axis a satellite would not stay there without active corrections. Periodical manoeuvres are needed. In the example treated here a manoeuvre is done every two weeks. A manoeuvre consists of a small change in velocity along the Earth-Moon-axis as it can be induced by a cold gas thruster [7].

The mechanics of Libration points are very complicated and determining the manoeuvres needed to stay there as well. Calculating this would go beyond the scope of this bachelor thesis. This is why the values of the needed velocity changes were determined by trial and error. Therefore, this needs to be seen as a basic investigation on the feasibility of such a mission rather than a detailed analysis.

To make biweekly manoeuvres possible they need to be of an accuracy of at least $1\frac{\mathrm{mm}}{\mathrm{s}}$. This is clearly feasible as cold gas thrusters can produce thrusts of down to 20mN [7]. With a satellite mass of 500kg this would result in a thrust time of 25s. This can be calculated using the conservation of linear momentum:

$$m \cdot v = \int F dt = F \cdot t \qquad \Rightarrow \qquad t = \frac{m \cdot v}{F} = \frac{500\mathrm{kg} \cdot 1\frac{\mathrm{mm}}{\mathrm{s}}}{20\mathrm{mN}} = 25\mathrm{s} \tag{4.12}$$

Constant coverage cone

As the satellite is moving on a non-circular orbit around the libration point L2 azimuth and elevation under which it is seen from a spots on Moon's surface vary over time. For spots near the edge of the covered area this can also result in loosing the connection. To calculate the cone in which a constant connection is guaranteed the worst-case must be assumed. In terms of elevation and range relative to the center of Moon's far side the worst-case is the combination of the lowest elevation and the shortest range. These values can be taken from the simulation. An applicable plot of elevation and range over time is shown in Figure 4.30.

The half apex angle of the constant covered cone can be calculated by using the laws of cosine and sine. Figure 4.31 shows the geometry of the problem.

The minimal radius $R_{S,min}$ of the satellite can be calculated out of Moon's radius R_M, the minimum distance $r_{S,min}$ to the centre of Moon's far side C and the minimum elevation $E_{S,min}$ under which the satellite is seen from that point by using the law of cosine:

$$R_{S,min}^2 = R_M^2 + r_{S,min}^2 - 2 \cdot R_M \cdot r_{S,min} \cdot \cos(90° + E_{S,min}) \tag{4.13}$$

Using this value the angle $\beta_{S,max}$ between the satellite radius vector and the vector pointing from M to C can be calculated with the law of sine:

$$\frac{\sin(90° + E_{S,min})}{R_{S,min}} = \frac{sin(\beta_{S,max})}{r_{S,min}} \Rightarrow \beta_{S,max} = \arcsin\left(\frac{r_{S,min}}{R_{S,min}} \cdot \cos(E_{S,min})\right) \tag{4.14}$$

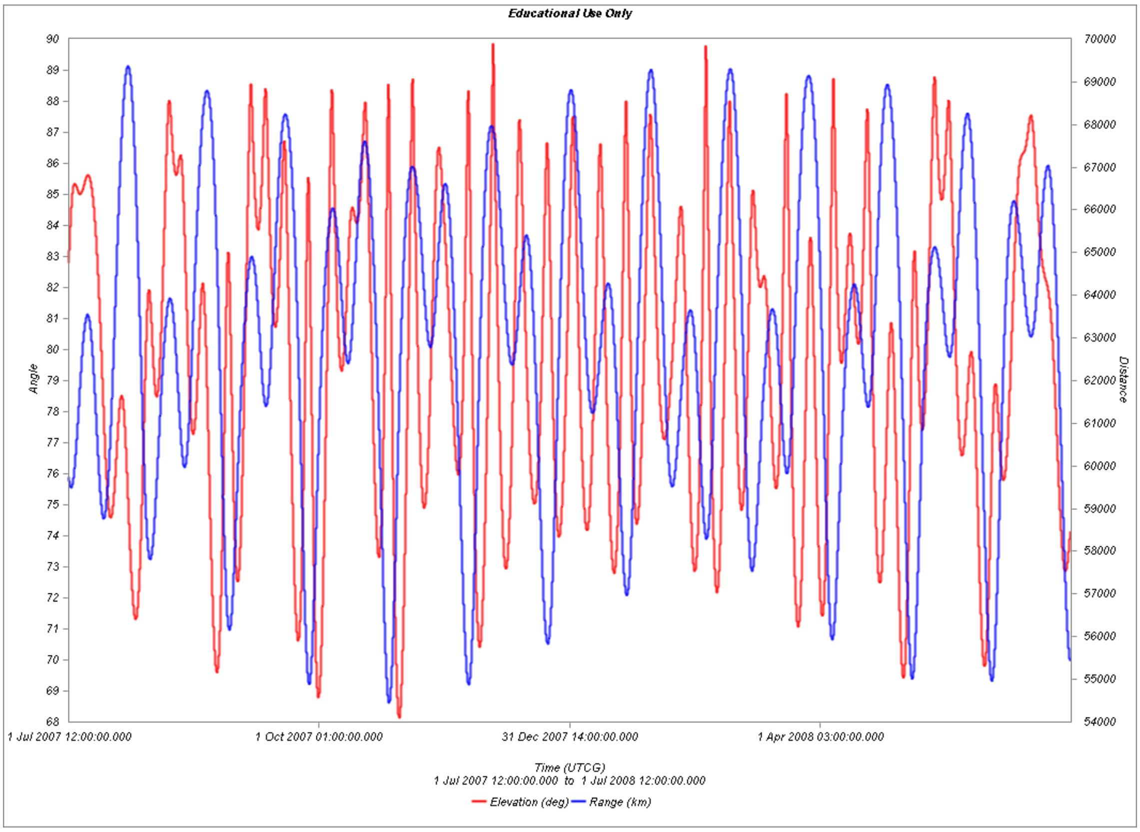


Figure 4.30: Elevation and range over time relative to the centre of Moon's far side. Source: STK

The last angle needed to calculate β_{max} is the apex angle γ of the triangle AMS at the point S. It can be calculated out of $R_{S,min}$, R_M and the minimum elevation E_0 under which a communication with the satellite is possible, using the law of sine:

$$\frac{\sin(90° + E_0)}{R_{S,min}} = \frac{sin(\gamma)}{R_M} \quad \Rightarrow \quad \gamma = \arcsin\left(\frac{R_M}{R_{S,min}} \cdot \cos(E_0)\right) \tag{4.15}$$

With a simple angle summation the half apex angle β_{max} of the constant covered cone can be obtained:

$$\beta_{max} = 180° - 90° - E_0 - \gamma - \beta_{S,max} \tag{4.16}$$

With a minimum elevation of the satellite seen from the centre of Moon's far side $E_{S,min} = 68°$ and a minimum range of $r_{S,min} = 54000\text{km}$ obtained from Figure 4.31, the radius of the Moon $R_M = 1737.1\text{km}$ [8] and a minimum elevation under which communication is possible of $E_0 = 5°$, the half apex angle of the constant covered cone is:

$$\beta_{max} \approx 62°$$

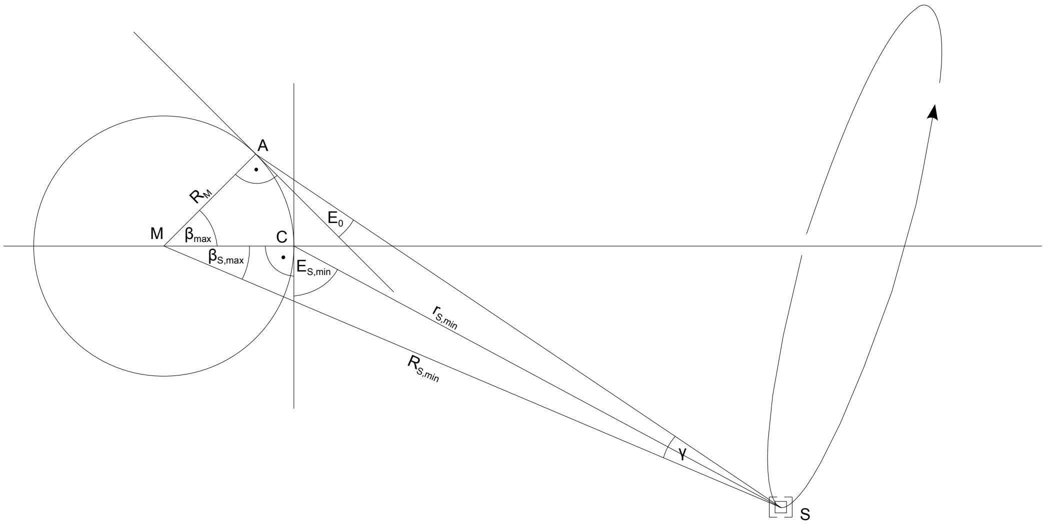


Figure 4.31: Geometry of the coverage cone calculation. On the left side is the Moon and to the right the satellite (S) on it's orbit around the L2. The centre of Moon's far side is located in C and the rim of the coverage cone is represented by the point A

Minimum height

The L2 relay satellite can not only be used by objects on ground but also by other satellites orbiting the Moon. This has the benefit that a constant connection to earth is possible even while the satellite is behind the Moon. But as the L2 relay satellite is not covering the whole far side of the Moon, the satellites orbiting the Moon need a minimum height to always be able to establish a link to Earth either directly or via the relay satellite.

The calculation is very similar to the coverage cone above. Figure 4.32 shows the geometry of the problem.

The calculations in the triangle AMS are now easier as it is rectangular. The new angle β'_{max} can be calculated out of Moon's radius R_M and the minimum radius of the L2 relay satellite $R_{S,min}$ as calculated above in the "coverage cone" paragraph by using a simple cosine:

$$\beta'_{max} = \arccos\left(\frac{R_M}{R_{S,min}}\right) - \beta_{S,max} \tag{4.17}$$

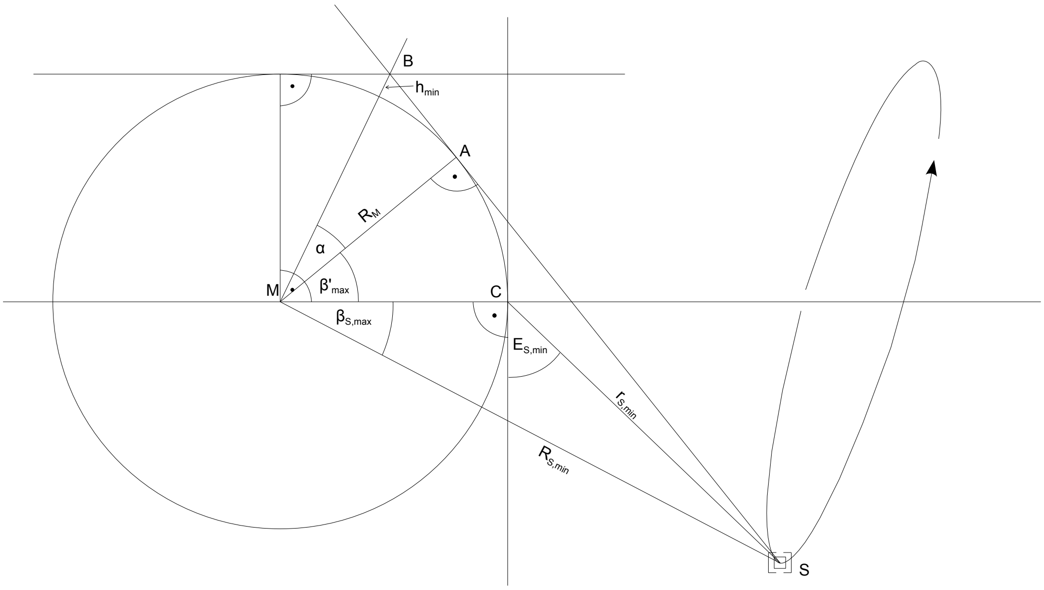


Figure 4.32: Geometry of the minimum height calculation. The Moon with it's centre M is located on the left side and the L2 satellite to the right. Point B is the intersection of the coverage edges of the connection to earth and to the L2 satellite. Satellites that are higher above the ground as point B always have a connection to Earth.

Using a simple angle summation the angle α can be determined:

$$\alpha = \frac{1}{2}(90^\circ - \beta'_{max}) \tag{4.18}$$

With this angle α and another cosine the minimum orbit height needed for a constant connection to Earth can be calculated:

$$h_{min} = \frac{R_M}{\cos(\alpha)} - R_M = R_M\left(\frac{1}{\cos(\alpha)} - 1\right) \tag{4.19}$$

With the same values for $E_{S,min}$, $r_{S,min}$ and R_M as in the coverage cone calculation above, the minimum orbit height needed for a constant connection either directly or via the L2 relay satellite is:

$$h_{min} \approx 27\text{km}$$

This is as already expected from the geometry in Figure 4.32 very small in comparison to common orbit heights. Some margin needs to be added but nevertheless, most missions would be able to establish a constant connection to earth using the L2 relay satellite.

Conclusion

The L2 relay satellite is a very good solution to cover Moon's far side, but with higher operational efforts than other options. The biggest benefit is that only one satellite is needed to provide a nearly constant link to Earth. It's weak point is the edge region. A constant coverage on ground can not be guaranteed here.

For satellite missions not lower than approximately 50km altitude a constant link to earth can be realised. This can be very interesting for missions observing the far side of the Moon with instruments producing huge amounts of data as the need for storing the results in a mass memory disappears.

5 Protocols

In classic space missions links to and from the satellites are direct, using a dedicated ground station. The ground stations are rented or owned by the satellite operator and the whole communication is scheduled to a specific time. New commands are prepared far before their execution so that there are multiple contact chances left if one ore more transmissions fail. In other words the high effort needed for operating a satellite mission mainly is due to the high effort that is put into communication with the satellite. Even more effort is required if relay satellites are needed. This is mainly the case in rover missions, as rovers normally do not have the transmission power to send radio waves all the way back to earth or do not even see the earth.

The Consultative Committee for Space Data Systems (CCSDS), an interagency working group defining standards for space applications, founded 1982 made a lot of progress in establishing standards for space communication links [1]. It is mainly thanks to them, that ground stations have standard interfaces to the satellites and mission control centres nowadays. Multiple agencies can work together on operating a mission without defining the communication protocols every time and a lot of the communication on ground is already handled by the internet. But the used communication protocols for space links are still very unintelligent.

When designing a satellite communication network for satellites and other objects in space, the protocols must be more flexible and self organised. Otherwise the operational effort would be excessive. The internet protocols Transmission Control Protocol (TCP) and Internet Protocol (IP) are the successful solution in the internet. So why not consider to use them for a space application? The flexible routing and the possibility to easily expand the network of IP as well as the fault tolerance and reliability of TCP are considerable attributes for a satellite communication network.

However, some characteristics of the space environment will cause problems. The higher bit error rates for example will cause TCP's slow start mechanism to throttle down the transmission speed lower than the link capacity. This is an issue that needs to be solved.

Other problems are the link delay and the fact that satellites might need to store packets until the next possibility to send it to the destination. A solution for these problems is currently under development by a group around Vint Cerf on behalf of NASA. It is called Delay Tolerant Networking (DTN) and can also be combined with TCP/IP [4].

Both solutions shall be covered here. As well as different communication scenarios at the Moon and their requirements on the protocols.

5.1 TCP/IP

5.1.1 IP

The Internet Protocol (IP) is the protocol defining the internet. In the Open Systems Interconnection Reference Model (OSI Model) it has the function of layer three, the Network Layer (see Figure 5.1) and thus has the following tasks:

1. ”The Network Layer provides the functional and procedural means for connectionless-mode or connection-mode transmission among transport-entities and, therefore, provides to the transport-entities independence of routing and relay considerations.

2. The Network Layer provides the means to establish, maintain, and terminate network-connections between open systems containing communicating application-entities and the functional and procedural means to exchange network-service-data-units between transport-entities over network-connections.

3. It provides to the transport-entities independence from routing and relay consideration associated with the establishment and operation of a given network-connection. This includes the case where several subnetworks are used in tandem [...] or in parallel. It makes invisible to transport-entities how underlying resources such as data-link-connections are used to provide network-connections.

4. Any relay functions and hop-by-hop service enhancement protocols used to support the network-service between the OSI end systems are operating below the Transport Layer, i.e. within the Network Layer or below.” [6]

In other words, it shall handle the network and all the routing and be transparent to higher layers.

The internet is designed hierarchically. A backbone network spans it worldwide and is divided into subnetworks locally. These subnetworks are then again divided into sub-subnetworks and so on. The address under which a client can be reached, called the IP address, contains the information to which subnetwork it belongs. This makes it possible to find the destination of a packet very fast. Routers check in which subnetwork the destination is and then pass the packet to the responsible router. This router then checks in which subsubnetwork the destination is and passes the packet accordingly and so on. A packet from one client to another client in a completely different subnetwork thus makes all the way up in the hierarchy and down again to the destination.

This makes it very useful for a space communication network, status and availability information just needs to be broadcasted locally. A satellite at the Moon does not need to know how the router configuration on earth is. It just knows that the packets destination is on earth because of the address and sends it to the next node which is mentioned as router to Earth in it's routing table. In terms of subnetworks the network at the Moon would be a subnetwork for it's own and the ground stations on earth it's routers to the earth internet.

However there is a difference to Earth subnetworks. Earth subnetworks are normally static and a satellite communication network is not. In a static subnetwork one node is the router to other subnetworks and this doesn't change over time. Thus the routing tables of the clients are fix. In satellite networks however the nodes move relative to each other which causes constant changes of the available links. The routes must constantly be adapted to the current situation. But this is no problem for the IP, the satellites just need to be aware of the current situation and adapt their routing tables accordingly.

5.1.2 TCP

To make communication reliable is the main purpose of the Transmission Control Protocol (TCP) in the internet. In the OSI model it is attached to layer four, the transport layer (see Figure 5.1) which has the following tasks:

1. ”The transport-service provides transparent transfer of data between session-entities and relieves them from any concern with the detailed way in which reliable and cost effective transfer of data is achieved.

2. The Transport Layer optimises the use of the available network-service to provide the performance required by each session-entity at minimum cost. This optimisation is achieved within the constraints imposed by the overall demands of all concurrent session-entities and the overall quality and capacity of the network-service available to the Transport Layer.

3. All protocols defined in the Transport Layer have end-to-end significance, where the ends are defined as transport entities having transport associations. Therefore, the Transport Layer is OSI end open system oriented and transport-protocols operate only between OSI end open systems.

4. The Transport Layer is relieved of any concern with routing and relaying since the network-service provides data transfer from any transport-entity to any other, including the case of tandem subnetworks [...]” [6]

In other words, the Transport Layer shall establish an end-to-end connection between two nodes and make the communication between them reliable and cost effective so that protocols above (next layer is Session Layer) do not need to deal with network resources.

To fulfil these requirements TCP has congestion and flow control and error detection mechanisms.

The error detection mechanism is based on checksums and acknowledgements. Every packet contains a checksum the receiver can validate it's correctness with. If the packet is received and correct, the receiver sends an acknowledgement to the sender. If the sender does not receive an acknowledgement in time it transmits the packet again.

The flow control mechanism ensures, that the receiver buffer will not exceed. Therefor the remaining buffer size is sent by the receiver in every acknowledgement. The sender stops to send when the receiver buffer is full and waits for a message indicating that the receiver is ready to receive data again.

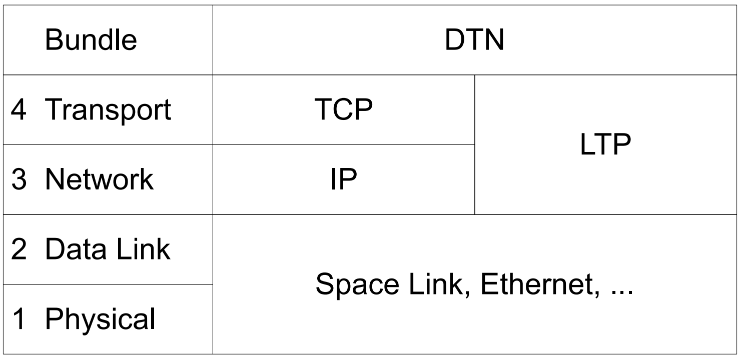
Finally the congestion control mechanism takes care that the network will not become overloaded. This is done by testing out the capacity of the connection. Therefor the transmission is started with the so called slow start. The congestion window that defines the number of not acknowledged packages that is allowed on the link, is set to a very low value and increased every time a positive acknowledgement is received. When a packet is lost the sender enters slow start again.

This works out perfectly for wired networks, where packet losses mainly occur when routers are overloaded and start to drop packets because of full buffers. For wireless networks this behaviour causes the transmission to be much slower than the link's ability. On a wireless link lost packets are much more likely but most times not a reason for a needed slow start. The main reason for package loss is the unreliable transmission medium. This is a problem for space links as well as for Wifi networks. One approach to solve it is to introduce a fast retransmit, if a packet is lost. The slow start phase is only entered after several lost packets. This is done in TCP-Reno.

However for severe round-trip delays TCP is even with adaptions ineligible as the protocol is very much based on close statement-answer loops with the receiver. To reach the link capacity with the transmission speed for example takes multiple round-trips. With round-trip times of several minutes this is very inefficient.

5.2 DTN

Delay Tolerant Networking (DTN) has a different attempt to solve the problems with high delays and bit error rates on long distance radio links. DTN introduces a new layer above the Transport Layer called Bundle Layer (see Figure 5.1) with the purpose to handle multiple hops with waiting times in the nodes and unite networks and links with different underlying protocols. Thus local TCP/IP networks on Earth and Moon or other planets can be interconnected by a protocol more suitable for a long distance radio link. And even if this link is not always available applications do not need to take care of scheduling their communication as the bundle layer does this.



<table>
<tr><td>Bundle</td><td colspan="2">DTN</td></tr>
<tr><td>4 Transport</td><td>TCP</td><td rowspan="2">LTP</td></tr>
<tr><td>3 Network</td><td>IP</td></tr>
<tr><td>2 Data Link</td><td colspan="2" rowspan="2">Space Link, Ethernet, ...</td></tr>
<tr><td>1 Physical</td></tr>
</table>

Figure 5.1: OSI model protocol stack showing TCP, IP and LTP with Bundle Layer extension including DTN.

But at first some deeper introductions into what bundles are and what the Bundle Layer does. A bundle contains a huge amount of data in the ideal case everything of one transmission. If for example multiple correlated files shall be downloaded from a satellite, all these files should be packed in one bundle. This makes it possible for the Bundle Layer to optimise the routing as fragmentation of transmission can be avoided. The benefit of this mechanism can be best shown with an example. If there is a link from a router to the destination, or a router near it and this link will terminate soon but multiple bundles still wait for transmission and not all of them can be transmitted, the bundle layer can decide to send multiple small bundles instead of starting the transmission of one large one, which could not be transmitted completely anyway. The link can be used much more efficiently this way, as the fragment of the large bundle would not be of any use for the receiver. It needs to wait for the next possible link anyway.

This leads directly to the next mechanism of the Bundle Layer. The possibility to store bundles in routing nodes until the next hop is available. If for example a satellite is relaying data from a ground vehicle on the far side of the Moon to Earth and is not positioned in an orbit around the Libration Point L2 (see Chapter 4.3.2), it needs to store the data first as it has no link to Earth at the moment itself. Hence the bundle is first transmitted to the relay satellite which stores it until it raises out of Moon's shadow and forwards the bundle to Earth.

The Bundle Layer expects the underlying protocols to be capable of transferring data reliably which eliminates the need for acknowledgements on bundle level. However they are still a possible option. This requirement for reliable communication makes TCP an excellent candidate for local networks. But as described in Chapter 5.1.2 it has some disadvantages on very long range links such as to Mars, which are disqualifying it. A new approach that is especially designed for DTN is called Licklider Transmission Protocol (LTP) which is named in honour of the American computer scientist Joseph Carl Robnett Licklider. The LTP approach to deal with severe round-trip times is to send large blocks and not to wait for an acknowledgement to continue with the next block. It distinguishes between red blocks which need to be transmitted reliably, like file headers, and green blocks where bit errors are acceptable, like the pixels of an image. When transmitting a red block, the receiver answers with an acknowledgement indicating which parts were read correctly and the sender can then retransmit missing parts.

5.3 Communication scenarios

5.3.1 Scenario 1: real-time

The simplest scenario is with a satellite as information source and a constant communication link to earth. Either directly or via a relay like the L2 relay satellite from Chapter 4.3.2. The router on the source satellite needs to be aware of it's and other relaying nodes orbits. This scenario can be realised with IP and a constantly updated routing table.

5.3.2 Scenario 2: real-time with intercepted connection

If interceptions of the connections are taken into account the situation becomes more complex and can not be handled by simple routing table updates. This scenario occurs if a relay satellite is needed but not available. For example a satellite behind the Moon and without a L2 relay satellite from Chapter 4.3.2, or a rover on the near side with a too low powered radio transmitter to establish a link to Earth.

Three states are now possible: direction, relayed and no communication. The orbit aware router does now need to buffer data during times of no connection. This can be realised by extending the buffer sizes of IP so that all the packets send during phases of no connection can be stored until the next possibility to forward them. The other solution is using DTN, which is designed for such situations, for storing and routing.

5.3.3 Scenario 3: non-real-time

What has not be taken into account until now is the case of a rover on the far side of the Moon where a Moon orbiting relay satellite needs to store the data the rover wants to send to Earth. This scenario is not real-time anymore as the receiving time is delayed in comparison to the sending time. However the solution for this scenario is very similar to the real-time with intercepted connection scenario. The only difference is that the relay satellites also need extended buffers now. This can again be realised with IP or DTN.

5.3.4 Scenario 4: optimisation of non-real-time scenario

In the previous scenarios the situation was very much simplified as the problems caused by multiple relay satellites were not considered. In a relay satellite constellation however multiple possible relays are the case. The sender somehow needs to decide which relay to take.

The easiest routing algorithm is to take the first available satellite always. A better solution would be to take the satellite that gets in contact with Earth again next, but this requires additional knowledge of all the satellites orbits.

5.3.5 Scenario 5: finite buffer sizes, multi-user case with random access

To finally get to the complexity a real satellite communication is of, multiple users must be taken into account. This introduces new problems like contention of communication with the relay and contention of memory on the relay satellite which leads to a more complex situation on the satellite.

For routing solutions to satisfy this problem additional information of the other users are mandatory. Different levels of detail of this information are applicable:

- The number of upcoming connections and their mean data volume is known. In this scenario two possible solutions for optimising the routing strategy are conceivable. Every user optimises it's routing for it's own delay or the mean delay.
- Reservations are possible and known by all users as well as the buffer levels of the relays. The relay satellites calculate a buffer overflow likelihood based on their current buffer levels, buffer sizes, reservations, density of sources and remaining time in shadow and distribute it to all the users. The users can decide then which satellite to use depending on the buffer overflow likelihood and delays the relay satellites bring along.
- The third possibility is that all participating nodes know all influencing factors and can calculate the optimum routing strategy. This however needs a lot of broadcasting and calculation power in the nodes.

5.4 Conclusion

For a communication network at the Moon DTN is not mandatory. TCP and IP will most probably do a good job. The round trip delay to Moon is still small in comparison to other Planets. It's maximum is:

$$t_{rt} = \frac{2 \cdot r_{M,max}}{c} = \frac{2 \cdot 405,500\mathrm{km}}{299,792\frac{\mathrm{km}}{\mathrm{s}}} \approx 2.7\mathrm{s} \tag{5.1}$$

It is very likely that TCP is still able to handle this delay and DTN with LTP is not necessarily needed. When thinking of a future satellite network at Mars or planets even farther away the use of a protocol like DTN is obligatory. Because of this the test in a communication network at the Moon is very interesting. The routing however is also in DTN still a field of research and was only touched in this thesis.

6 Outlook

The most interesting aspects of a satellite communication system were attended to or touched in this bachelor thesis. However multiple topics for further research are still available. In the field of the Libration Point orbit for example a lot of work is still needed to be done to actually implement such a mission. Examining and planing the mission to a Libration Point is very demanding.

More detailed investigations are also needed on the protocols. Simulations especially of the various routing strategies will most probably bring up new conclusions and new problems.

The motivation for this thesis was the communication experiment LunaNet which will be a payload of the European Student Moon Orbiter. This mission will be a great possibility to test technologies of the space internet in the near future. The focus is hereby to test the protocols and also routing strategies can be tested up to a certain level. The work on finding a proper orbit constellation for satellite communication systems done in this thesis will not have an impact on the LunaNet experiment as the orbit is defined by other factors. The research on the protocols however will continue as a part of this mission.

Bibliography

[1] CCSDS. CCSDS History. Website, 2009. Available online at http://public.ccsds.org/about/history.aspx visited on October 27th 2009.

[2] Institute for Communications and Navigation at TUM. Satellite Navigation Laboratory lecture notes, 2008.

[3] Network Working Group. RFC 2581: TCP Congestion Control. Website, 1999. Available online at http://tools.ietf.org/html/rfc2581 visited on October 28th 2009.

[4] Network Working Group. RFC 4838: Delay-Tolerant Networking Architecture. Website, 2007. Available online at http://tools.ietf.org/html/rfc4838 visited on October 28th 2009.

[5] Network Working Group. RFC 5325: Licklider Transmission Protocol - Motivation. Website, 2007. Available online at http://tools.ietf.org/html/rfc5325 visited on October 28th 2009.

[6] ITU-T. ITU-T Recommendation X.200: Information Technology - Open Systems Interconnection - Basic Reference Model: The Basic Model. Website, 1994. Available online at http://www.itu.int/rec/T-REC-X.200-199407-I/en visited on October 28th 2009.

[7] Wilfried Ley, Klaus Wittmann, and Willi Hallmann. *Handbuch der Raumfahrttechnik*. Hanser, 2008.

[8] Dr. David R. Williams (NASA). Moon Fact Sheet. Website, 2006. Available online at http://nssdc.gsfc.nasa.gov/planetary/factsheet/moonfact.html visited on October 28th 2009.

[9] Information Sciences Institute University of Southern California. RFC 791: Internet Protocol. Website, 1981. Available online at http://tools.ietf.org/html/rfc791 visited on October 28th 2009.

[10] Information Sciences Institute University of Southern California. RFC 793: Transmission Control Protocol. Website, 1981. Available online at http://tools.ietf.org/html/rfc793 visited on October 28th 2009.