

Stirling Engine Manual

© Franco Normani

Email: franco@real-world-physics-problems.com

Alternate email: fvnorman@hotmail.com

March 09, 2013

Table of Contents

Legal Notice and Disclaimer	Page 4
References	Page 4
Some Background.....	Page 5
So What Is The Purpose of this Stirling Engine Program?.....	Page 6
Stirling Engine Basics – What is a Stirling Engine?	Page 7
The Three Stirling Engine Configurations.....	Page 9
Outline.....	Page 12
The Program Input	Page 13
Input File Parameters and What They Mean	Page 15
Using the Program	Page 37
Method of Solution	Page 38
A Closer Look at the Heater and Cooler.....	Page 41
A Closer Look at the Regenerator	Page 42
Graph to help you visualize the temperature of the working gas as it flows through the heater, regenerator, and cooler.....	Page 46
Engine Losses	Page 47
A Closer Look at Power and Efficiency	Page 48
Buffer Space.....	Page 51
A Closer Look At The Displacer	Page 53
Flywheel Design	Page 55
Miscellaneous Items.....	Page 61
The Radiator and Heat Source	Page 66
Offset Angle in the Expansion and Compression Space Volume.....	Page 66
Excerpts Taken From Stirling Engine Design Manual	Page 67
Sensitivity Analysis	Page 69
Analyzing the Program Results	Page 78
Summary of units used.....	Page 88
Troubleshooting	Page 89
Technical Support.....	Page 89

Stirling Engine Thermodynamics	Page 90
The Heater, Cooler, and Non-Ideal Regenerator	Page 109
Energy Balance for the System.....	Page 115
Regarding Engine Pressure	Page 116
A Closer Look at Power Calculation	Page 116
Numerical Solution Method.....	Page 118
Appendix A – Crank Drive Design For Alpha Engines	Page 120
Appendix B – Ross Yoke Design For Alpha Engines	Page 122
Appendix C – Crank Drive Design for Beta and Gamma Engines	Page 125
Appendix D – Carnot Efficiency	Page 126
Appendix E – MOD II Engine.....	Page 128
Appendix F – SES Stirling Dishes.....	Page 133
Appendix G – Example Optimization	Page 136

Legal Notice and Disclaimer

This program and manual are for personal use only. You are not permitted to sell or redistribute this program and manual in any way.

This program is intended to aid designers and enthusiasts in building Stirling engines for their own home and personal use. The results of the program are accurate as far as the physics and mathematics are concerned. But you are still expected to exercise good judgment when designing and building a Stirling engine (or any engine). Therefore, I'm not responsible for the use or misuse of the program, or the information presented here. If you plan on building an engine for commercial application it is strongly advised that you seek consultation from an expert who can give you professional advice based on their experience with commercial systems.

References

[1] "Nonlinear Analysis of Stirling Engine Thermodynamics", Oak Ridge National Laboratory, R. D. Banduric, N. C. J. Chen, June 1984.

Web link:

http://www.ornl.gov/sci/ees/etsd/btrc/eere_research_reports/thermally_activated_technologies/engine_driven/stirling_rankine/modeling_and_simulation/ornl_con_154/ornl_con_154.pdf

[2] "Fundamentals of Heat and Mass Transfer", Fourth Edition, School of Mechanical Engineering, Purdue University, Frank P. Incropera, David P. DeWitt, 1996. Page 424-425.

[3] "Effects of Pressure-Drop Correlations On Stirling Engine Predicted Performance", Oak Ridge National Laboratory, N. C. J. Chen, E. P. Griffin, August 1983.

Web link:

http://www.ornl.gov/sci/ees/etsd/btrc/eere_research_reports/thermally_activated_technologies/engine_driven/stirling_rankine/modeling_and_simulation/ornl_conf_830812_52/ornl_conf_830812_52.pdf

[4] "Engineering Thermodynamics", Second Edition, Stanford University, University of Arizona, Department of Mechanical Engineering, W. C. Reynolds, H. C. Perkins, 1977.

[5] "Numerical Study on Optimal Stirling Engine Regenerator Matrix Designs Taking Into Account The Effects of Matrix Temperature Oscillations", *Energy Conversion And Management*, volume 47, pages 894-908, Technical University of Denmark, S. K. Andersen, H. Carlsen, P. G. Thomsen, June 2005.

Some Background

The Stirling engine has attracted much attention over the years. Its potential for high efficiency and the ability to use a wide variety of fuels has made it a serious contender for alternative power sources, especially in automotive applications. The potential for Stirling engines to replace the Internal Combustion engine (ICE) in automobiles was explored in the late 1970's and 1980's.

In 1986 a technical report was released by NASA outlining the development of the MOD I, and MOD II automotive Stirling engine (see Appendix E for details and pictures). The engine used pressurized hydrogen as the working gas. It was developed and produced by a collaborative effort between NASA and MTI (Mechanical Technology Incorporated). The MOD II engine in particular could reach a thermal efficiency of 38.5% (significantly higher than a spark ignition ICE), and with power comparable to an ICE of the same size (83.5 hp). It burned fuel with cleaner emissions than an ICE due to the fact that it burns fuel *externally* to the engine. It also produced much less noise during operation. Therefore, no muffler or catalytic converter would be needed for the tailpipe.

The technical problems were for the most part related to the high-pressure requirements of the engine (up to 15 MPa), which requires bulky components and specialized seals. Nevertheless it was projected that the cost of production due to economics of scale would be competitive with an ICE. Unfortunately, it failed to attract large investment in the automotive industry mainly because it was still in its infancy development-wise, and it couldn't compete with an ICE on the basis of responsiveness. Stirling engines do not respond as quickly to changes in power requirements as an ICE and take longer to warm up before reaching full power. Great strides could be made in solving this dilemma by using elaborate control systems regulating pressure, etc. But in the words of the large automakers such as GM, the market "would not tolerate this", even though efficiency gains and reduction in harmful tailpipe emissions were significant. Furthermore, cheap energy (oil) was abundant and no one at the time cared to invest heavily in an engine deemed "economically risky".

And now here we are today, faced with an energy future which can no longer blindly rely on fossil fuels to satisfy our needs. Conservation is becoming more important as well as alternative sources of energy. The Stirling engine is one of the possibilities for weaning ourselves off fossil fuels, given that it can use any source of heat to run, including biomass and solar. It is becoming clear that interest in this type of engine isn't going to go away. In fact, there is growing interest. One only has to look at <http://www.stirlingenergy.com> and <http://www.stirlingbiopower.com> to see that there are already some major strides being made. Appendix F shows pictures of Stirling engines used in solar power applications, manufactured by Stirling Energy Systems (SES).

Currently, there are many researchers, designers and enthusiasts who aspire to build Stirling engines. A search on *Google* will result in many "hits" showing the level of interest out there.

So What Is The Purpose of this Stirling Engine Program?

The purpose of this program and manual is to help the enthusiast, designer or inventor design a working Stirling engine without getting bogged down by the mathematics and thermodynamics.

Stirling engine design and optimization is not a trivial matter. At its core, it's an area that requires an understanding of thermodynamics, fluid mechanics, heat transfer, and material science. This takes time and effort to become familiar with and is surely the reason that most people get discouraged. From this point of view it would then seem that Stirling engine design is more involved than ICE design, even though it requires none of the complex mechanisms such as valves and timing mechanisms. But in reality it's just a different kind of engine.

Nevertheless, people tend *not* to have an intuitive feel for how to make a Stirling engine. The problem is that good design (optimization) requires the balancing of several conflicting variables, and it can be difficult to keep them straight. The solution to this is often experimentation, which unfortunately can be time and resource intensive. The alternative is to model the physics of the engine as accurately as possible, and reduce the trial and error in the construction stage, as a result.

The program I created (in Fortran) is a numerical model of the physics of a Stirling engine, based on consideration of the thermodynamics, fluid mechanics, and heat transfer principles involved. This model is simplified only where it's very difficult to do otherwise, but has enough detail in it to provide reasonable calculations on engine performance.

One does not need to directly understand the physics and mathematics involved to use the program. One only needs to input the design dimensions, and change certain parameters until optimization is reached. A detailed description of the various parameters and their physical significance will be given as you read on.

This program and manual is the result of a year and a half of learning about Stirling engines; both from reading lots of information found online, such as research papers, and from my own design efforts.

Admittedly, even though I have a background in mechanical engineering I have found it challenging to make sense of a lot of what I read. One reason for this is because I believe that the richness in information available is somewhat diminished by less than clear or contradictory explanations. One of the reasons for this, as I see it, is that Stirling engine theory (given its inherent physical complexity) is many times reduced to simplifications, which are *incomplete*. So it is unavoidable that various incomplete theories will not mesh well together, and this will tend to create confusion. To address this I personally feel that the top-down approach works best. It is much better to make an effort to understand the *entire* system first, and then figure out which approximations to use and why. Doing this

also lets you better understand where the incomplete solutions are coming from and in what context they apply.

Another reason why technical explanations found in the literature (and online sources) are so often difficult to understand is because they are written by people who seem more interested in communicating what they know in a manner most accessible to those who are already familiar with the subject. For various reasons, they are either discouraged or are oblivious towards expressing the information in a clear manner understandable to those who are *not* familiar with the subject. With this in mind there are a lot of papers out there that are essentially transcripts of the “internal dialogue” of authors intimately familiar with different subjects. The unfortunate result is a lot of quality information that is obscured with technical jargon and arcane language, making it difficult for outsiders to grasp. This can create the wrong impression (on both sides) that it’s “too hard” and not worth the effort unless you’re part of an elite group. As a result, the flow of knowledge is restricted. This is true of most presentations of scientific knowledge, covering a wide range of topics (not just Stirling engines).

So I created this program and manual in order to help bridge this gap and help a novice make sense of Stirling engines and their design, using straightforward language and clear explanations. Clarity was the only criterion I followed when writing this program and manual. I did not try to adhere to a particular format or style just for the sake of doing so. In academic circles this would no doubt cause a stir. But who cares. As long as *you* can understand it, is all that matters.

My intent was to put together a resource that is accessible to someone who wants a laypersons understanding, and which (hopefully) will not be a struggle to understand if they are not part of a certain “culture”. That said, this is a laypersons guide, but it will still take some effort for you to understand given the nature of the subject. But you won’t be turned off by obscure and overly technical language.

In large part, this manual is based on the references given above. During my searches I found these references to be particularly useful and informative. They will be cited throughout this manual.

Stirling Engine Basics – What is a Stirling Engine?

A Stirling engine is a heat engine that works on the basis of an external applied temperature difference. By maintaining a hot and cold temperature difference the engine is able to run and produce mechanical power. It is different from the Internal Combustion Engine (ICE) in that it is a closed cycle; that is, the working gas is enclosed (sealed) inside the engine. This is in contrast to the ICE in which the working gas (air) is drawn in from the environment, combusted with fuel, and expelled as exhaust. In such an engine valves and timing mechanisms are necessary. But in a Stirling engine, no such components are required. In addition, the Stirling engine is not restricted to the type of fuel used. It is indifferent to the source of heat, which opens up many possibilities,

including non-polluting solar energy, or the burning of biomass (wood, husks, ethanol, etc), which are carbon-neutral. Carbon-neutral means they absorb as much carbon dioxide (during their growth - due to photosynthesis) as they emit when burned. This is unlike fossil fuels, which add a net amount of carbon dioxide to the atmosphere when burned.

The basic principle of the Stirling engine is this. The engine is filled (under pressure), with a gas such as air, helium, or hydrogen. This is called the “working gas”. Inside the engine the gas is heated. This increases its pressure and moves pistons as a result. The gas is then cooled, lowering its pressure. It is then heated again, and the cycle repeats. In a real engine this typically happens very fast, on the same order of speed as an ICE. The working gas is shuttled back and forth very quickly inside the engine, between the hot and cold ends, continuously gaining and losing heat and producing power as a result.

The working gas inside the engine is heated with a heater, and cooled with a cooler.

The heater and cooler are typically compact heat exchangers consisting of narrow tubes (or passageways) in which the working gas flows. It is through these passageways that the working gas either gains heat (becoming hotter), or loses heat (becoming cooler).

The outside surface of the heater is exposed to a source of high temperature, such as the flame of a burner, or concentrated solar energy. The outside surface of the cooler is exposed to a source of cold temperature such as ambient air, or water.

In between the heater and cooler is a regenerator. A regenerator increases the efficiency of a Stirling engine by lowering the heat input requirement of the heater and the heat removal requirement of the cooler. It is not necessary to have a regenerator for the engine to run but in the interest of cost-reduction, especially where the cost of heater fuel is concerned, it is wise to have one.

The way the regenerator works is by storing some of the heat energy of the working gas as it moves from the heater to the cooler, thereby reducing the cooling demand on the cooler. And on the return path, as the working gas moves from the cooler to the heater, it “gains” back some of that heat energy, thereby reducing the heating requirement of the heater. A regenerator basically pre-heats the working gas before it enters the heater, and pre-cools the working gas before it enters the cooler.

The regenerator is usually made of an intricate matrix material, made of stacked metal screens or metal felt, woven from fine wire. This provides the large surface area necessary for efficient heat exchange with the working gas.

In general, when designing Stirling engines for high power and efficiency there are several main factors, which must be addressed:

(1) Keep dead volume to a minimum. Dead volume decreases engine power. Dead volume is the volume that is “unswept” by the motions of the pistons. This is the volume

contained in the heater, cooler, regenerator, and all the clearance spaces. This volume is constant at all times.

(2) Design the heater to maximize heating of the working gas, i.e. once the gas exits the heater its temperature must be as close as possible to that of the heater walls. This can be accomplished by using narrow and long tubes/passageways for the gas to flow through.

(3) Design the cooler to maximize cooling of the working gas, i.e. once the gas exits the cooler its temperature must be as close as possible to that of the cooler walls. This can be accomplished by using narrow and long tubes/passageways for the gas to flow through.

(4) Design the regenerator to maximize heat exchange with the working gas. This can be accomplished by using a sufficiently dense matrix material with large surface area.

(5) Keep pumping losses to a minimum. Pumping losses are the friction (flow) losses caused by the working gas as it “pushes” through the narrow tubes/passageways of the heater and cooler, and the regenerator matrix. Minimizing these losses can be accomplished by using a large number of tubes/passageways in the heater and cooler, and a regenerator with a large volume.

Now, points (2) - (5) can all be satisfied at the same time. However, satisfying these points is in direct conflict with point number (1). If one wishes to keep dead volume as small as possible they will have difficulty designing a good heater, cooler, and regenerator while keeping pumping losses low.

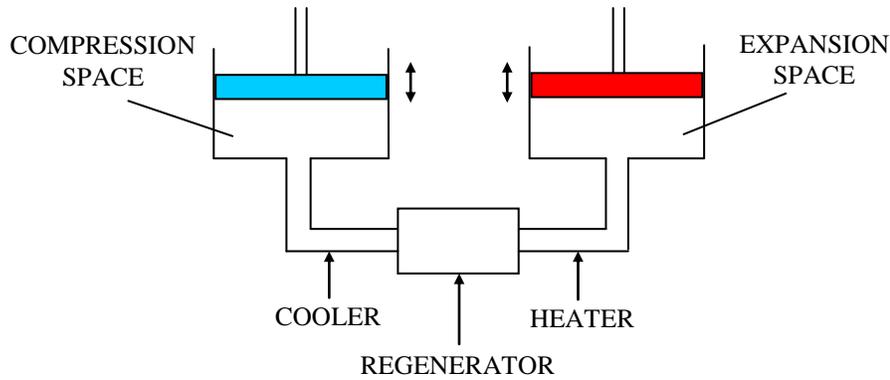
The solution is to compromise somewhat on all five points in order to achieve the best design possible. The challenge is to find the optimal balance resulting in the best engine design. This is what the program does.

The Three Stirling Engine Configurations

There are three standard configurations for Stirling engines. They are: ‘alpha’, ‘beta’, and ‘gamma’ engines.

Alpha Engine

The figure below shows a standard alpha engine.

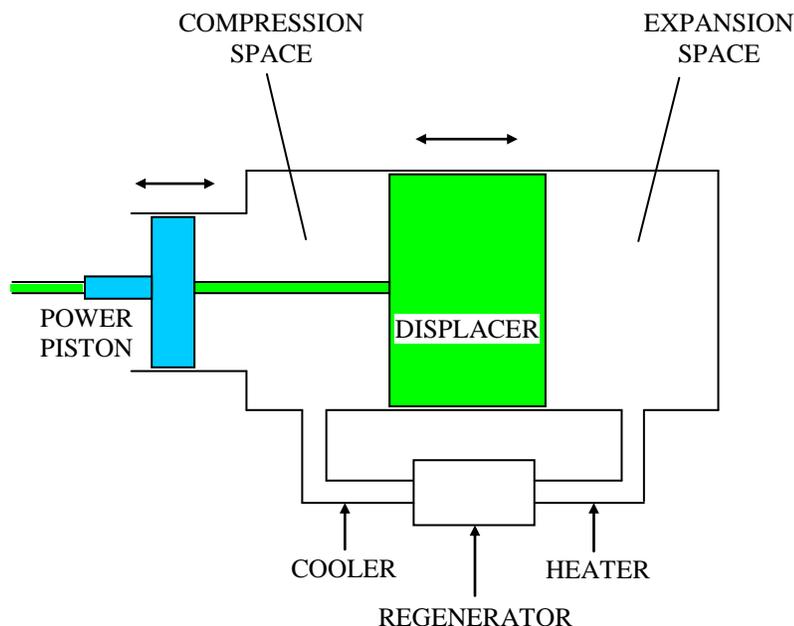


The working gas inside the engine is repeatedly “shuttled” back and forth between the expansion and compression space, due to the up-and-down motion of the two pistons. This repeatedly forces the working gas back and forth through the heater, regenerator and cooler. As a result, the gas is repeatedly heated and cooled, and power is produced.

Alpha engines are the simplest to understand, and are the easiest to construct. It’s also easy to minimize the dead (clearance) volume in the expansion and compression spaces. But one of their main disadvantages is that they can require temperature resistant seals for the piston exposed to high temperature (shown on the right). The seals for the other piston (on the left) don’t have to be temperature resistant because it is constantly exposed to cool temperatures, due to its physical proximity to the compression space.

Beta Engine

The figure below shows a standard beta engine.



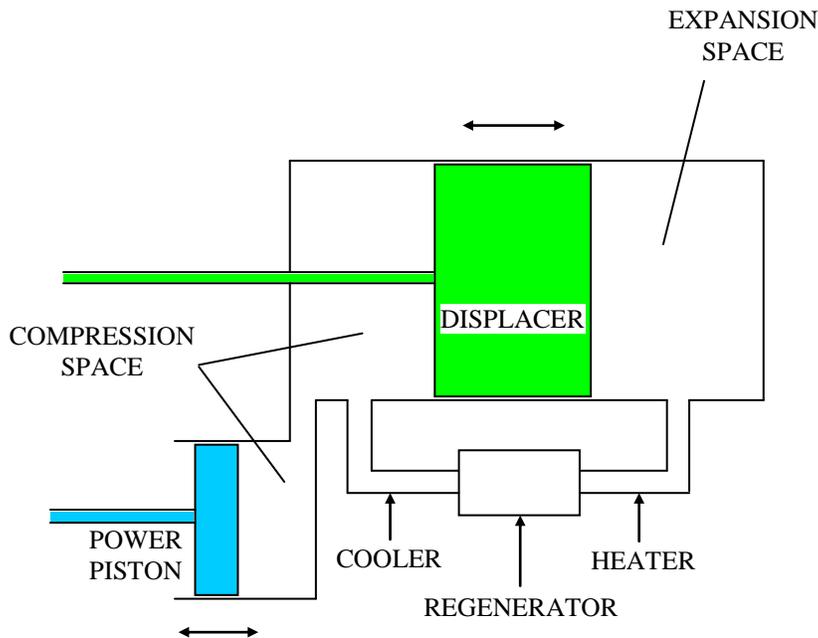
Beta engines are compact in size. They use a power piston and displacer, which are in line with each other. Unlike an alpha engine that uses two pistons, the beta engine uses one power piston and displacer. The purpose of the power piston is to generate power, while the purpose of the displacer is to move the working gas back and forth through the heater, regenerator, and cooler. As a result, the pushing force experienced by the displacer is very little compared to that of the power piston.

The power piston, displacer, and displacer rod are sealed around their gaps to prevent the leakage of working gas. The seal for the displacer is placed on the end closest to the compression space, in order to avoid direct contact with the hot working gas (in the expansion space). As a result, this seal does not need to be temperature resistant. The seals for the displacer rod and power piston do not need to be temperature resistant either since they are constantly exposed to cool engine temperatures. This is due to their physical proximity to the compression space.

A disadvantage of the beta engine is that it can be difficult to minimize the dead (clearance) volume in the expansion and compression space, given that there must be enough clearance to allow the working gas to “feed in” unobstructed from the heater and cooler.

Gamma Engine

The figure below shows a standard gamma engine.



Gamma engines are the same as beta engines, except that the power piston is “shifted” down. This can make it easier to construct the mechanical drive and linkages since the power piston and displacer are a certain distance apart (instead of aligned with each other). For this reason, the gamma configuration is often the preferred choice by Stirling engine enthusiasts.

A disadvantage of the gamma engine is that it unavoidably introduces dead volume in the compression space due to the physical separation of the displacer and power piston.

Outline

As mentioned, this program and manual is intended to help you optimize your Stirling engine design and reduce trial and error in the building stage. However, this is not a design manual in the sense that it does *not* tell you how to join metals, what type of materials to use and their dimensions, or how to manufacture the different parts. A basic explanation of some of these things will be given, but for the most part the information presented will mostly serve to help you decide what the “best” design is. For example, you will find out how much power and efficiency you can roughly expect from your design. You will find out the speed at which maximum power occurs. You will find out how big and heavy to make the flywheel in order to ensure “smooth” engine operation.

Furthermore, the program will allow you to optimize your design so that you get the highest power possible. In particular, this program will allow you to determine the approximate size of regenerator, and the number of heater and cooler tubes/passageways necessary for maximum power. This optimization will be based on the swept volumes and other design parameters that you input into the program. This will all be described in full detail in the following sections.

The first sections of this manual describe how to use the program, and also provide useful insight into Stirling engines and their design. These sections will be marked as *must-read*.

The later sections will explain more of the theory used in the model. They will outline the mathematics and physics of the program. Some background in fluid mechanics and thermodynamics will be necessary to make sense of it. However, it is not necessary to fully understand or even read these sections in order to use the program. These sections will be marked as *optional*, meaning you don’t have to read them.

Note that the program only applies to dual-piston alpha engines or single piston/single displacer beta and gamma engines (as will be shown below). No provision is made in the program to account for multi-piston and multi-cylinder engine configurations.

In addition, this program only applies to “kinematic” engines, where the piston and displacer motions are prescribed; meaning they are connected to linkages and

crankshaft. There are other types of Stirling engines known as Free-piston engines, which do not require linkages and crankshaft. The program does not model these.

The Program Input (must-read)

All of the input required for the program to run is entered in a single input file, called 'input'. A sample of this input file is shown below (in smaller text):

```
Swept volume in expansion space (cm^3)
87.0
Swept volume in compression space (cm^3)
87.0
Minimum volume in expansion space (cm^3)
5.0
Minimum volume in compression space (cm^3)
5.0
Phase angle (degrees)
90.0
Regenerator void volume (cm^3)
72.0
Flow cross-sectional area of each heater tube (cm^2)
0.04
Flow cross-sectional area of each cooler tube (cm^2)
0.04
Wetted perimeter for each heater tube (cm)
0.8
Wetted perimeter for each cooler tube (cm)
0.8
Flow length of each heater tube (cm)
18.0
Flow length of each cooler tube (cm)
18.0
Number of tubes in heater
50
Number of tubes in cooler
50
Heater temperature (K)
1000.0
Cooler temperature (K)
293.0
Average pressure (Pa)
8785750.0
Gas constant (J/(kg.K))
287.0
Cp of working gas (J/(kg.K))
1007.0
Cv of working gas (J/(kg.K))
720.0
Viscosity (N.s/m^2)
0.00003
Regenerator constant eps1 (dimensionless)
0.0
```

Regenerator constant eps2 (dimensionless)

0.0

Scale factor to get the average pressure you want at the engine speed you want (default = 1.0)

1.0

As you can see, different lines of text describe the input, and the input numbers are located directly below these lines of text. You simply change the numbers (do not alter the lines of text) to match your desired input. You then save the file, and run the program.

The input file can be opened with the Notepad editor. To do this, click the file with your right mouse button. Click on “Open”, click on Notepad, and then click OK to open the file. Make sure “Word Wrap” (in the Format menu) is not selected.

To save the file (after making changes), click on “File” at the top left of the menu, and then click “Save”.

Make sure you only edit and save the file in the Notepad editor. There is no guarantee that the program will run error-free if you use an editor such as WordPad, or Microsoft Word. Also, make sure you don’t have an extension at the end of the filename, such as “.txt”. Simply leave the name as ‘input’ (not ‘input.txt’), otherwise it won’t run.

If you accidentally erase a line in the input file, or alter the positioning of anything such as a line of text, or a number, you may copy and paste the above sample input file into Notepad, and then save it as-is. This way if you accidentally corrupt your input file, it can easily be replaced. If viewing this manual with Adobe Reader, you can copy the input file as follows: First, click on “Tools” at the top menu, then click “Basic”, then click “Select”. This will allow you to highlight the text you want. In your case highlight the sample input file by holding down the left mouse button and then “drag” over the file until it is all selected. Then press Ctrl-C (hold down Ctrl and press C). This copies the text. Next, go to a blank Notepad editor and press Ctrl-V (hold down Ctrl and press V). This will “paste” the input file into Notepad. You can now save the file as ‘input’. When you save the file the first time it will unavoidably be saved as ‘input.txt’. You have to remove the “.txt” at the end. This can be done manually (after saving) using the Rename function. To use the Rename function click the file with your right mouse button and then click “Rename”. This will allow you to rename the file. But in your case all you will do is delete “.txt” at the end.

In addition:

- Make sure there are no blank lines above the first line, which is: “Swept volume in expansion space (cm³)”.
- The positioning of the lines of text and the numbers in the input file must be preserved; otherwise you run the risk of the program not running properly. That said it’s always a good idea to make a backup of the original input file, just in

case. This way you don't have to bother with copying and pasting the sample input file into Notepad (as shown above), if anything goes wrong.

- Only use integers (e.g. 45, 68) for the number of tubes. For all other values express as decimal (e.g. 140.0, not 140; 0.0, not 0).

For this next part I'm going to go through the various lines of the input file and explain what they all mean. To make it easier to read I'm going to make the lines of the input file bold and with larger font. And underneath I'm going to explain what they mean.

Input File Parameters and What They Mean (must-read)

First group of input file parameters:

Swept volume in expansion space (cm³)

87.0

Swept volume in compression space (cm³)

87.0

Minimum volume in expansion space (cm³)

5.0

Minimum volume in compression space (cm³)

5.0

Phase angle (degrees)

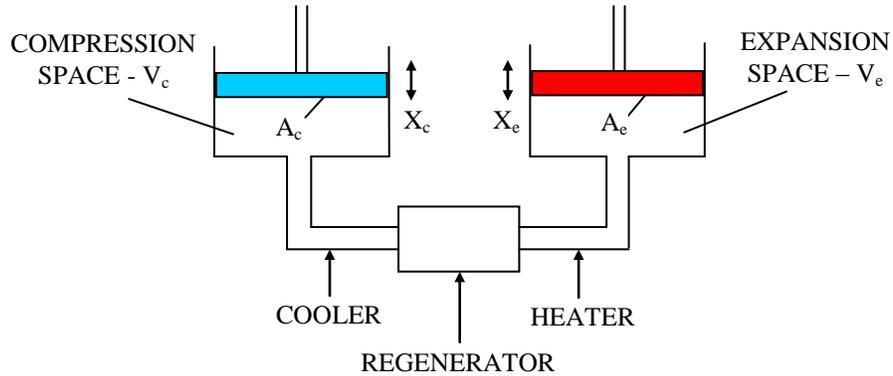
90.0

(Note: cm³ means "cubic centimeters")

First, it is necessary to show what the expansion and compression space is (and introduce some additional variables as well). I will show this for the three types of Stirling engine configurations: 'Alpha', 'Beta', and 'Gamma'. These configurations are shown in the figures below.

For all three configurations the heater feeds into the expansion space, and the cooler feeds into the compression space.

STIRLING ENGINE WITH 'ALPHA' CONFIGURATION



Where:

A_e = the area of the expansion space piston (this is the cross-sectional area of the face)

A_c = the area of the compression space piston (this is the cross-sectional area of the face)

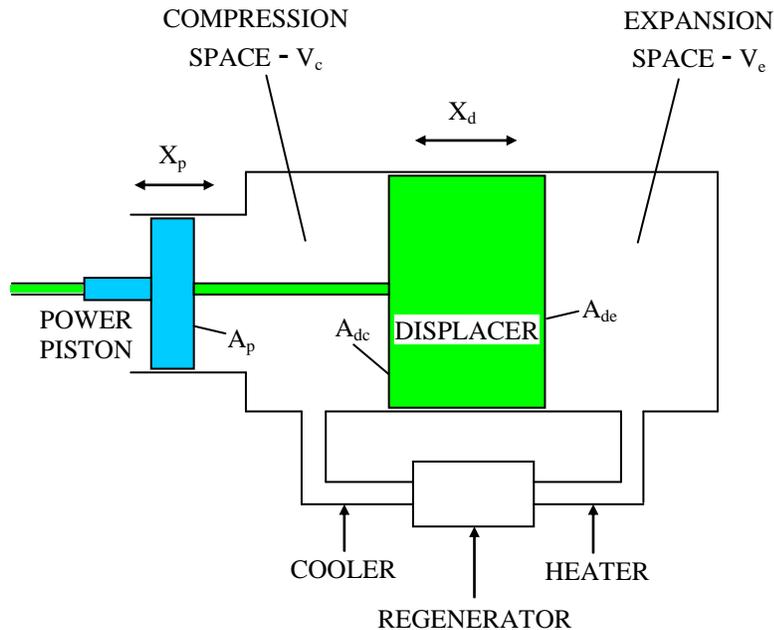
X_e = the stroke length of the expansion space piston

X_c = the stroke length of the compression space piston

V_e = the volume in the expansion space

V_c = the volume in the compression space

STIRLING ENGINE WITH 'BETA' CONFIGURATION



Where:

A_{de} = the area of the displacer on the expansion space side

A_{dc} = the area of the displacer on the compression space side. Note that this area is A_{de} minus the area of the displacer rod.

A_p = the area of the power piston

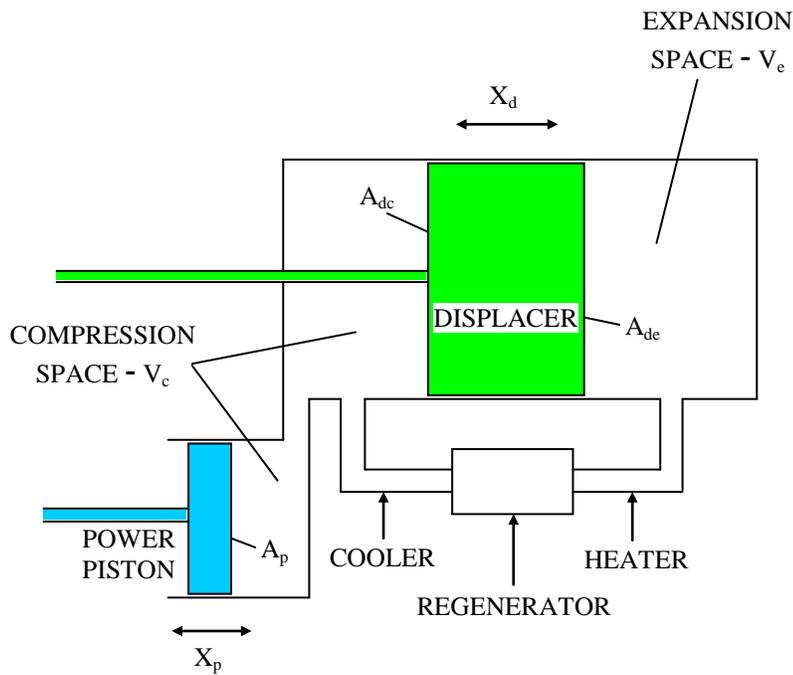
X_d = the stroke length of the displacer

X_p = the stroke length of the power piston

V_e = the volume in the expansion space

V_c = the volume in the compression space

STIRLING ENGINE WITH 'GAMMA' CONFIGURATION



Where:

A_{de} = the area of the displacer on the expansion space side

A_{dc} = the area of the displacer on the compression space side. Note that this area is A_{de} minus the area of the displacer rod.

A_p = the area of the power piston

X_d = the stroke length of the displacer

X_p = the stroke length of the power piston

V_e = the volume in the expansion space

V_c = the volume in the compression space

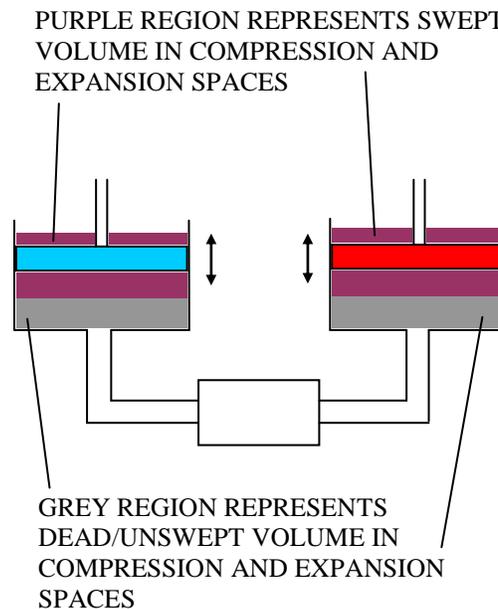
Note that the above variables are exactly the same as for a beta engine.

Notice that a gamma engine is simply a beta engine with the power piston “shifted” down. This results in a simpler construction for the drive mechanism given that the displacer and power piston are physically offset from each other.

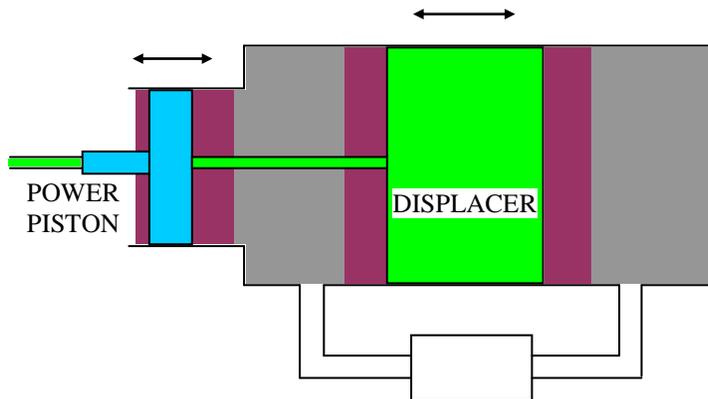
For a beta engine the power piston is usually the same diameter as the displacer, and they are both inside the same cylinder. In the figure above the power piston and displacer are of different diameter, and are in two separate cylinders. This was drawn just to show the general case.

Next, I will illustrate the concept of “swept” and “unswept” volume. This will be shown for the three engine configurations, in the figures below.

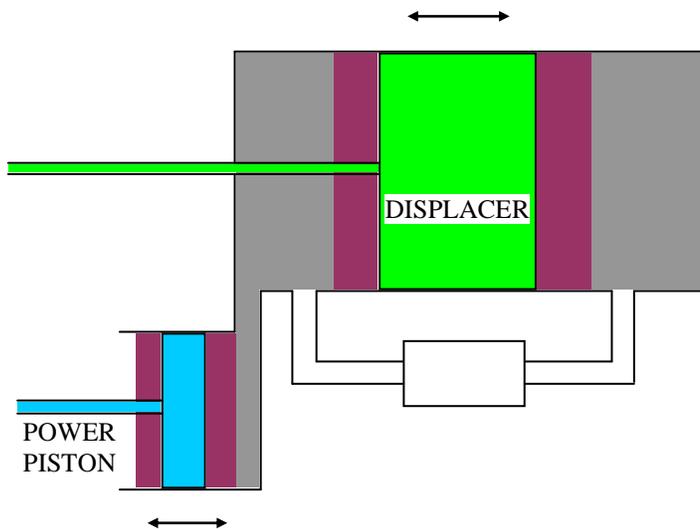
STIRLING ENGINE WITH ‘ALPHA’ CONFIGURATION



STIRLING ENGINE WITH 'BETA' CONFIGURATION



STIRLING ENGINE WITH 'GAMMA' CONFIGURATION



As you can see, the swept (purple) regions in the expansion and compression space are the volumes that are “swept” by the motion of the pistons and displacers. And the dead/unswept (grey) regions in the expansion and compression space are the volumes that remain “untouched”, or “dead”. These regions are “untouched” by the displacer and the power piston, during their motion.

For example, in an alpha engine, if the expansion space piston diameter is 10 cm and the stroke length is 8 cm, the swept volume is:

$$V_{swept} = \frac{\pi}{4} 10^2 \times 8 = 628.3 \text{ cm}^3$$

The ratio between piston diameter and stroke length is typically 1:1 to 2:1. But the exact ratio used will depend on the details of your design. However, the thermodynamics of the model is indifferent to this ratio. But in your design this ratio may become important for practical reasons.

For the three engine configurations the instantaneous volume (at each point in time) in the expansion and compression space can easily be calculated. In this program a sinusoidal variation for volume is assumed. This is not necessarily the case, but it is a good approximation that helps to simplify the calculations.

The mathematical expressions for volume variation are given below. It is not necessary to know them but they do help you to understand the values that go into the input file. To make sense of these expressions, it's helpful to refer to the figures above.

Alpha Engine

In the expansion space the volume is:

$$V_e = V_{\min,e} + \frac{1}{2} \cdot V_{swept,e} + \frac{1}{2} \cdot V_{swept,e} \cdot \sin(\theta + \beta)$$

In the compression space the volume is:

$$V_c = V_{\min,c} + \frac{1}{2} \cdot V_{swept,c} + \frac{1}{2} \cdot V_{swept,c} \cdot \sin(\theta)$$

where

$$V_{swept,e} = A_e X_e$$

and

$$V_{swept,c} = A_c X_c$$

A_e = the area of the expansion space piston

A_c = the area of the compression space piston

X_e = the stroke length of the expansion space piston

X_c = the stroke length of the compression space piston

$V_{\text{swept},e}$ = the swept volume in the expansion space (this is shown as the purple region in the expansion space). This value goes into the input file.

$V_{\text{swept},c}$ = the swept volume in the compression space (this is shown as the purple region in the compression space). This value goes into the input file.

$V_{\text{min},e}$ = the minimum volume in the expansion space. This is also called the clearance volume. For an alpha engine, this volume is equal to the dead/unswept volume in the expansion space (this is shown as the grey region in the expansion space). This value goes into the input file.

$V_{\text{min},c}$ = the minimum volume in the compression space. This is also called the clearance volume. For an alpha engine, this volume is equal to the dead/unswept volume in the compression space (this is shown as the grey region in the compression space). This value goes into the input file.

θ = the crank angle of rotation (this changes with time)

β = the phase angle (a constant) between the expansion and compression space volume. This value goes into the input file.

Physically speaking, in an alpha engine, the expansion space piston is “ahead” of the compression space piston by an angle β . For example, when the expansion space piston is at top-dead-position, the compression space piston is “behind” in rotation by an angle β .

Note that for alpha engines the maximum engine power and thermal efficiency occurs at $\beta \cong 90^\circ$ (this is discussed in more detail later).

There are two main drive mechanisms one can use for alpha engines. They are the Crank Drive and the Ross Yoke. The details of their design are given in Appendix A and B (page 120 and 122).

For beta and gamma engines the expression for volume in the expansion space is the same as for alpha engines. But the expression for volume in the compression space is more complex because it's related to the *combined* motion of the displacer and power piston. The easiest way to evaluate it is by expressing it mathematically and then graphing it alongside the expansion space volume. The graph can then be used to visually

determine the swept volumes, minimum volumes, and phase angle. These can then be entered directly in the input file. This will be explained next.

Beta and Gamma Engine

In the expansion space the volume is:

$$V_e = V_{\min,e} + \frac{1}{2} X_d A_{de} - \frac{1}{2} X_d A_{de} \cdot \sin(\theta + \alpha)$$

In the compression space the volume is:

$$V_c = V_{dead,c} + \frac{1}{2} X_p A_p + \frac{1}{2} X_d A_{dc} - \frac{1}{2} X_p A_p \cdot \sin(\theta) + \frac{1}{2} X_d A_{dc} \cdot \sin(\theta + \alpha)$$

Where:

A_{de} = the area of the displacer on the expansion space side

A_{dc} = the area of the displacer on the compression space side. Note that this area is A_{de} minus the area of the displacer rod.

A_p = the area of the power piston

X_d = the stroke length of the displacer

X_p = the stroke length of the power piston

$V_{\min,e}$ = the minimum volume in the expansion space. For a beta and gamma engine, this volume is equal to the dead/unswept volume in the expansion space (this is shown as the grey region in the expansion space).

$V_{dead,c}$ = the dead/unswept volume in the compression space (this is shown as the grey region in the compression space). For a beta and gamma engine, this volume is generally

not equal to the minimum volume in the compression space (due to the mathematical nature of the expression for V_c). The minimum volume in the compression space can be found by graphing V_c (shown below).

α = the phase angle (a constant) between the displacer and the power piston.

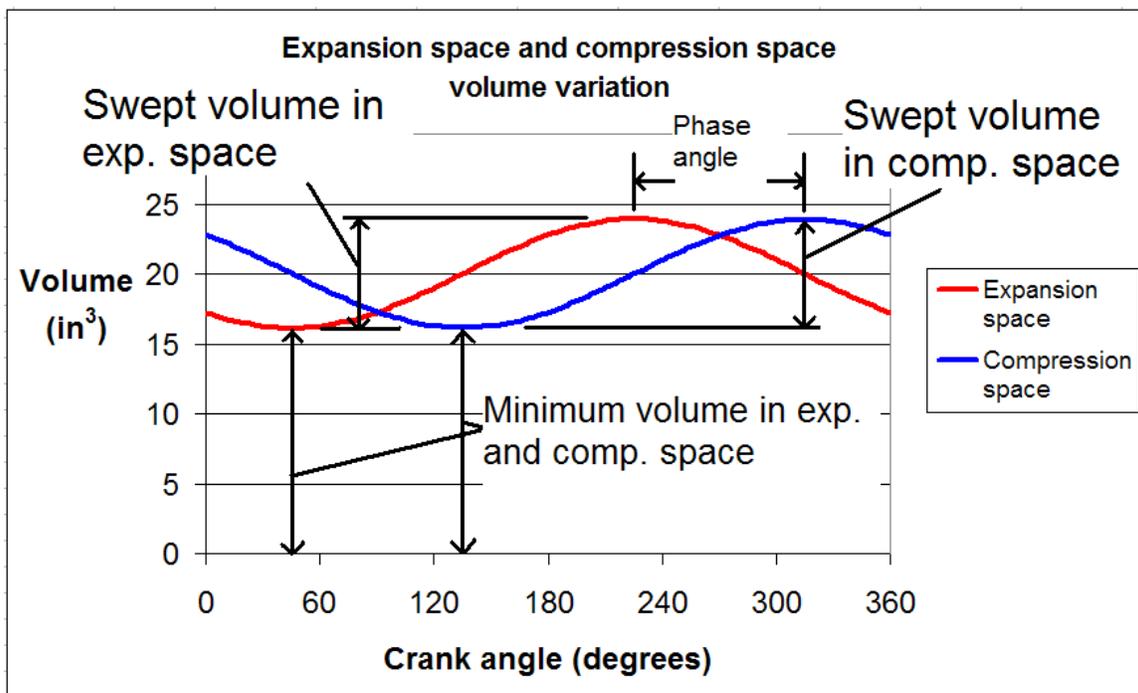
Physically speaking, in beta and gamma engines, the displacer is “ahead” of the power piston by an angle α .

Appendix C (page 125) describes a crank drive mechanism for beta and gamma engines.

For beta and gamma engines, the phase difference α that results in maximum power and thermal efficiency is the one that produces a phase difference of roughly 90° between the volumes V_e and V_c . This is the same as for alpha engines. This phase difference can be determined visually from graphing V_e and V_c (shown below).

The mathematical expressions for V_e and V_c (given on page 22) can now be inserted into a spreadsheet like Excel, and plotted graphically.

The graph below shows an example of the volume variations V_e and V_c as a function of crank angle θ . In this case I chose cubic-inches as the units for the volume. Obviously, you can use any units you like.



Note from the figure above that the curve for the expansion space volume is to the left of the curve for the compression space volume. This might seem counterintuitive given that the expansion space volume is “ahead” of the compression space volume, but if you think about it from a graphing perspective, you’ll realize that this is in fact the case.

Observing the figure above we can rewrite the volume variations in the same format as for an alpha engine:

$$V_e = V_{\min,e} + \frac{1}{2} \cdot V_{\text{swept},e} + \frac{1}{2} \cdot V_{\text{swept},e} \cdot \sin(\theta + \beta)$$

and

$$V_c = V_{\min,c} + \frac{1}{2} \cdot V_{\text{swept},c} + \frac{1}{2} \cdot V_{\text{swept},c} \cdot \sin(\theta)$$

Where:

$V_{\text{swept},e}$ = the swept volume in the expansion space (determined from the graph). This value goes into the input file.

$V_{\text{swept},c}$ = the swept volume in the compression space (determined from the graph). This value goes into the input file.

$V_{\min,e}$ = the minimum volume in the expansion space (determined from the graph). This value goes into the input file.

$V_{\min,c}$ = the minimum volume in the compression space (determined from the graph). This value goes into the input file.

β = the phase angle (determined from the graph). This value goes into the input file.

Using this graphical method one can obtain a mathematical expression for volume variation equivalent to the formulation for an alpha engine. This means that a beta and gamma engine can be reduced to an equivalent alpha engine.

The swept volumes and phase angle (shown in the graph) can be varied by adjusting the following values: α , A_{de} , A_{dc} , A_p , X_d , X_p . Ideally you would adjust these values until $\beta \cong 90^\circ$, knowing that this is the optimal phase angle.

To vary the minimum volume (in the expansion and compression space) one must adjust the dead/unswept volumes in these spaces (the grey regions shown in the figures).

Once again, note that in the expansion space, $V_{\min,e}$ is equal to the dead/unswept volume. So if $V_{\min,e} = 0$ this means that the dead/unswept volume in the expansion space is zero.

However, in a gamma engine $V_{\min,c} = 0$ is impossible due to the physical gap between the displacer and power piston. But for a beta engine it is possible to have $V_{\min,c} = 0$. This is possible if the swept volume of the displacer overlaps the swept volume of the power piston. But mathematically speaking, this means that $V_{\text{dead},c}$ must be a negative number. This may sound confusing, but it means that the individual swept volumes (purple regions) of the displacer and power piston must overlap. And this overlap volume is equal to $V_{\text{dead},c}$. And the largest possible (negative) value of $V_{\text{dead},c}$ is such that $V_{\min,c} = 0$. If it were any larger the displacer and power piston would “collide”.

However, to get maximum power, it is best to keep the minimum volume in the expansion and compression space as small as possible (for all three engine-configurations). In particular, for beta and gamma engines this can be achieved by having the heater tubes feed in through the top of the expansion space. This reduces the clearance requirement since the working gas isn't feeding in from the side. For the compression space, keeping minimum volume small usually means that there must be some overlapping swept volume between the displacer and power piston (negative $V_{\text{dead},c}$). But care must be taken to ensure there is adequate clearance during the part of the cycle where the working gas is entering the compression space (from the cooler). This occurs when the displacer is moving towards the expansion space (in other words, the expansion space is getting smaller).

On a final note, for all three engine-configurations the swept volume in the expansion and compression space is often the same but it doesn't have to be. In fact, the details of your design may dictate that they *should* be different.

Second group of input file parameters:

Regenerator void volume (cm³)

72.0

Flow cross-sectional area of each heater tube (cm²)

0.04

Flow cross-sectional area of each cooler tube (cm²)

0.04

Wetted perimeter for each heater tube (cm)

0.8

Wetted perimeter for each cooler tube (cm)

0.8

Flow length of each heater tube (cm)

18.0

Flow length of each cooler tube (cm)

18.0

Number of tubes in heater

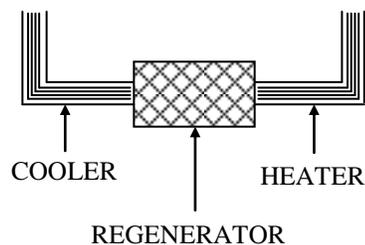
50

Number of tubes in cooler

50

(Note: cm^3 is “cubic centimeters”, cm^2 is “square centimeters”, and cm is “centimeters”)

These parameters deal with the specifics of the heater, regenerator and cooler. Let’s first show a schematic of the heater, regenerator, and cooler assembly, to help illustrate what they are.



Because the Stirling engine runs at high rotational speeds, the working gas oscillates rapidly back and forth inside the engine. As a result, there must be rapid heating and cooling of the gas as it moves from one part of the engine to another. This can only be accomplished by forcing the gas through heated, and cooled narrow tubes or passageways in order to achieve the high rate of heating and cooling required. However, the penalty for doing this is a so-called pumping (flow) loss, which results from forcing a gas through a narrow passage. Pumping losses will be discussed in greater detail later.

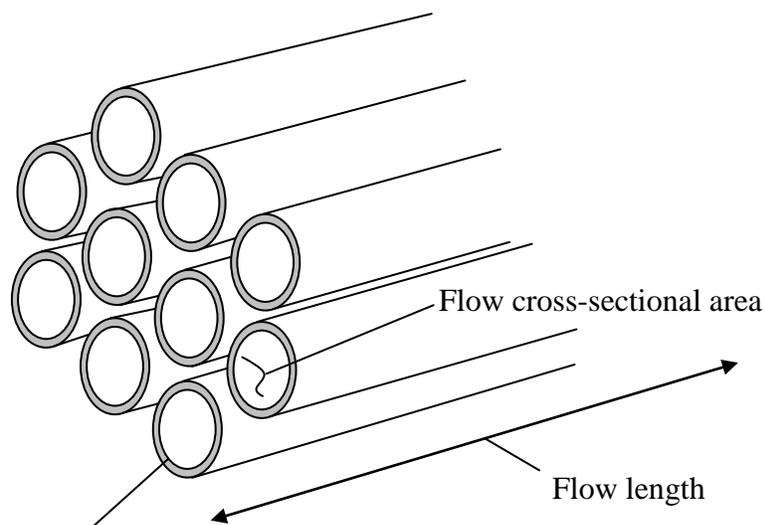
Stirling engine illustrations which show heating and cooling taking place at the cylinder wall of the expansion and compression space are generally not realistic. For engines of reasonable size and power the rate of heat transfer resulting from this would be far too small. However, for small model type Stirling engines (as seen on YouTube) one can heat and cool the walls of the expansion and compression space, and the engines will run at high speed. In these cases, because the engines are small, the heat transfer surface area is large compared to the volumes of the expansion and compression space. Therefore, the heating and cooling of the working gas is sufficiently “quick” to accommodate high engine speed.

In between the heater and cooler is a regenerator. A regenerator increases the efficiency of a Stirling engine. It is not necessary to have one for the engine to run but it does greatly reduce the energy input requirement of the heater and the energy removal requirement of the cooler, making the engine run more efficiently. The way the regenerator works is by storing some of the heat energy of the working gas as it moves from the heater to the cooler, thereby reducing the cooling demand placed on the cooler. And on the return path, as the working gas moves from the cooler to the heater, it “gains” back that heat energy, thereby reducing the heating demand on the heater.

A regenerator is typically made of a dense matrix consisting of stacked metal screens, or metal felt, woven from fine wires. This provides the high heat transfer capability needed to “store” and “release” heat.

The “regenerator void volume”, as specified in the input file, is simply the volume of empty space inside the regenerator housing. This is the space is not “filled” with matrix material. Therefore, the regenerator void volume is the *original* empty volume inside the regenerator housing minus the volume occupied by the matrix material. The specific value for the regenerator void volume will be discussed in greater detail below, following the description of the heater and cooler tubes (since they are related).

The input parameters: “flow cross-sectional area”, “wetted perimeter”, and “flow length” for each heater and cooler tube are illustrated in the figure below.



Wetted perimeter (the length around the inside - for a circular tube this is equal to the circumference around the inside)

The working gas flows through the inside of the tubes/passageways, gaining or losing heat depending on whether it's flowing through the tubes of the heater or cooler.

The “flow cross-sectional area” is simply the flow area on the inside of the tube. For a circular tube this is the area of a circle.

The “wetted perimeter” is the perimeter on the inside of the tube that is exposed to the working gas.

The flow length is the total length of the heater and cooler tubes.

The heater and cooler tubes can have different dimensions, depending on the requirements of your design.

As a rough guideline, the inside diameter of the heater and cooler tubes is around 2-5 mm. Therefore, the flow cross-sectional area is simply the area of a circle given the inside tube diameter.

Note that in some heater and cooler tubes, the passageways are not circular. They can be square or rectangular. In this case an effective (hydraulic) inside diameter D_h can be calculated:

$$D_h = \frac{4A_c}{P_w}$$

where A_c is the flow cross-sectional area for a tube, and P_w is the wetted perimeter.

In this case one would choose values for A_c and P_w so that D_h falls within the 2-5 mm range.

Note that the flow length of the tubes is not necessarily the same as the *heated* and *cooled* lengths. The heated length is the length of tube that is directly exposed to the hot source. The cooled length is the length of tube that is directly exposed to the cold source. An example of a hot source is a flame. An example of a cold source is a coolant, such as water.

However, the flow lengths that must be entered in the input file are the *total* tube lengths. The total length is the heated/cooled length *plus* any extra length at the ends, which inserts into the regenerator housing or the expansion/compression space cylinders. The amount of extra length will depend on the specifics of your particular design.

The length of tube that is directly exposed to the hot or cold source must be long enough to ensure sufficient heat transfer. In other words, the exposed length has to be long enough such that the working gas reaches a temperature close to that of the heater or

cooler tube wall temperature (possibly within a few degrees) by the time it exits the tubes. As it turns out, the required length can be approximated based on the inside diameter of the tube.

Assuming we have an efficient regenerator providing good heat exchange, the approximate ratio of exposed tube length to inside tube diameter is 80:1. For example, if the inside tube diameter is 1.5 mm, the exposed tube length should be about $1.5 \times 80 = 120$ mm.

This ratio of 80:1 is based on average ratios taken from the literature, on different Stirling engine designs. This ratio also matches well with theoretical heat transfer predictions. Page 71 of the following reference shows the ratios for different engines: “A Survey of Oscillating Flow in Stirling Engine Heat Exchangers”, T. W. Simon, J. R. Seume, University of Minnesota, March 1988. The web link is:

http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19880012938_1988012938.pdf.

This ratio of 80:1 holds well for high and low engine pressures, type of working gas, and engine speed.

However, if we were not using a regenerator, the tubes would require a much longer exposed length in order to reach a temperature close to that of the heater and cooler tube wall. Thus, the required ratio would be much larger than 80:1. This is because there is no regenerator to pre-heat and pre-cool the working gas before entering the heater and cooler tubes.

But in the program we are always assuming the use of a regenerator. The reason for this is simple. It is much more thermally efficient to operate an engine with a regenerator than without. However, it is possible to have higher engine power without a regenerator. This will be discussed in more detail later on.

Let's now calculate the volume of the heater and cooler, based on their exposed lengths. It is necessary to do this because the regenerator void volume is related to the volume of the heater and cooler, based on their exposed lengths.

The *heated* volume of the heater is given by:

$$V_{heater} = A_{heater} L_{heater} N_{heater}$$

Where:

A_{heater} = the flow cross-sectional area of each heater tube/passageway

L_{heater} = the length of each heater tube/passageway exposed to the hot source (this may be less than the total length).

N_{heater} = the number of heater tubes/passageways

The *cooled* volume of the cooler is given by:

$$V_{cooler} = A_{cooler} L_{cooler} N_{cooler}$$

Where:

A_{cooler} = the flow cross-sectional area of each cooler tube/passageway

L_{cooler} = the length of each cooler tube/passageway exposed to the cold source (this may be less than the total length).

N_{cooler} = the number of cooler tubes/passageways

Using the 80:1 ratio as an estimate, choose the tube dimensions such that $L_{heater}/(\text{inside-tube-diameter of heater tube}) = 80$, and $L_{cooler}/(\text{inside-tube-diameter of cooler tube}) = 80$.

Given that the heating and cooling requirements are generally similar (i.e. one must be balanced by the other), it makes sense to make the heater volume equal to the cooler volume (based on the exposed lengths). In other words $V_{heater} \approx V_{cooler}$. However, it is *not* necessary that they are exactly equal. Their exact value will depend on the details of your design.

The easiest way to make the heater and cooler volumes equal is to use equal flow areas, equal lengths, and an equal number of tubes for both. However, in your design it is possible that the heater tubes will be different in dimension than the cooler tubes. Therefore the numbers of tubes will be different in order to make the volumes equal.

The void volume of the regenerator can now be calculated.

The void volume of the regenerator is simply $V_{heater} + V_{cooler}$ (using the expressions given previously). In other words, the void volume of the regenerator is roughly twice that of either the heater or cooler (based on their exposed lengths), assuming they are both roughly the same volume ($V_{heater} \approx V_{cooler}$).

This value for regenerator void volume, relative to heater and cooler volume, is based on comparison of different Stirling engine designs in the literature. It was observed on several designs that the regenerator void volume is equal to the volume of the heater plus cooler (based on the exposed tube lengths), with a variability of around $\pm 40\text{-}50\%$. And it was also observed on several designs that the volumes of the heater and cooler are roughly equal, with a variability of around $\pm 25\%$.

So this begs the question: Is the best engine design one in which the heater volume equals the cooler volume, and the regenerator void volume is equal to the combined volume of the heater and cooler?

Not necessarily. In fact, there may be very good reasons why an optimal engine design will have significantly different design criteria than stated above. But the specifics of this will depend strongly on the details of the engine itself, which would no doubt involve careful heat transfer and flow loss considerations, as well as experimentation. But it is my feeling that any difference in design from what I outline here will *not* result in a significantly better engine with much higher power and efficiency. So as it stands, unless you know better, simply stick to the criteria I have given and you should be fine.

As mentioned before, the void volume of the regenerator is based on the *total* volume of the regenerator housing, and the volume of the matrix inside. So there needs to be a way of calculating void volume based on some parameter. The parameter commonly used is known as “porosity”. By definition, porosity is the fraction of empty volume with respect to total volume. For example, if the porosity of the regenerator is 80%, and the total volume inside the housing is 100 cm^3 , the void volume is $0.80 \times 100 = 80 \text{ cm}^3$.

One can also calculate the required volume of the regenerator housing knowing the porosity of the matrix and the regenerator void volume. For example, if the regenerator void volume is 75 cm^3 and the porosity of the matrix is 80%, the required (total) volume of the regenerator housing is $75/0.80 = 93.75 \text{ cm}^3$.

In a typical design the number of heater and cooler tubes can be anywhere from a few dozen to several hundred.

The main benefit of using a large number of tubes is that it “splits” up the flow and reduces pumping (flow) losses as a result.

For example, having two tubes would require much more pumping power than having ten tubes. But the downside of this is that you are adding dead volume to the engine and this tends to decrease engine power. Consequently, you have two opposing influences on engine power. Eventually, by increasing the number of tubes (and regenerator void volume, proportionally), maximum engine power is reached. The main purpose of the program is to find this point of maximum power.

As mentioned, the actual diameter and length of the tubes in your design will depend on what you have to work with. For example, it may depend on the size of your heat source. For a larger heat source area you will probably need longer tubes with a corresponding larger (inside) diameter, in order to match the 80:1 ratio. And for the cooler, depending on your cooling medium, you will need a certain length of tube and diameter as well. For example, if using a cooling medium such as water you can keep the tube length much shorter than if you are using, say, the surrounding air as the cooling medium. This is because water is much better at dissipating heat than air, so the “exposed” length can be much shorter. However, the surrounding air can still be used for cooling, by circulating

the water through a large radiator (exposed to the air). This cools the water for reuse in the cooler. This is an example of a good design.

Similarly, to enhance heat transfer you can place fins on the heater and cooler tubes. This also helps keep the exposed length as short as possible (which reduces dead volume).

Details such as these will depend on your design and on what you have to work with. It is intended that details such as these will be balanced along with the predictions of the program.

But keep in mind that the program will always output an answer. It is up to you to recognize that, in practice, the input must be realistic and fall within reasonable design guidelines such as ability to machine, produce within budget, material availability, etc.

Third group of input file parameters:

Heater temperature (K)

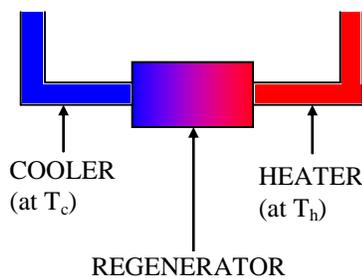
1000.0

Cooler temperature (K)

293.0

(Note: K is “Kelvin”, which by definition equals: (degrees Celsius) + 273.15)

From the figure below, the heater temperature (T_h) is the temperature of the hot source. The cooler temperature (T_c) is the temperature of the cold source.



Fourth group of input file parameters:

Average pressure (Pa)

8785750.0

(Note: Pa is “Pascals”)

This is the desired average pressure inside the engine. This pressure is based on the absolute pressure inside the engine. Mathematically, it is equal to the average pressure over a cycle.

Fifth group of input file parameters:

Gas constant (J/(kg.K))

287.0

Cp of working gas (J/(kg.K))

1007.0

Cv of working gas (J/(kg.K))

720.0

Viscosity (N.s/m²)

0.00003

These are constants based on the properties of the working gas. Note that it is not important to understand the units (J/kg.K) and (N.s/m²).

For convenience I will include the properties of some common gases here, given as nominal (average) values at lower pressure [2,4]. It is acceptable to use these as a decent approximation even though pressure inside Stirling engines can be very high (20-30 MPa (megapascals) for high power engines). But for lower power engines (say 1-2 horsepower), which may operate at an average pressure of 1-2 MPa, these constants may be a bit more accurate. For all the gases, except Carbon Dioxide, the properties are given at a temperature of about 350 degrees Celsius, which is the approximate average temperature in high-power Stirling engines. The properties for Carbon Dioxide are given at room temperature (25 degrees Celsius), since that is all I could find.

Air:

Gas constant = 287 J/kg·K

Cp = 1007 J/kg·K

Cv = 720 J/kg·K

Viscosity = 0.00003 N·s/m²

Argon:

Gas constant = 208 J/kg·K

Cp = 523 J/kg·K

Cv = 315 J/kg·K

Viscosity = 0.00004 N·s/m²

Carbon Dioxide:

Gas constant = 189 J/kg·K
Cp = 845 J/kg·K
Cv = 656 J/kg·K
Viscosity = 0.00002 N·s/m²

Helium:

Gas constant = 2077 J/kg·K
Cp = 5200 J/kg·K
Cv = 3123 J/kg·K
Viscosity = 0.00003 N·s/m²

Hydrogen:

Gas constant = 4120 J/kg·K
Cp = 14310 J/kg·K
Cv = 10190 J/kg·K
Viscosity = 0.000015 N·s/m²

Nitrogen:

Gas constant = 296 J/kg·K
Cp = 1038 J/kg·K
Cv = 742 J/kg·K
Viscosity = 0.00003 N·s/m²

Sixth group of input file parameters:

Regenerator constant eps1 (dimensionless)

0.0

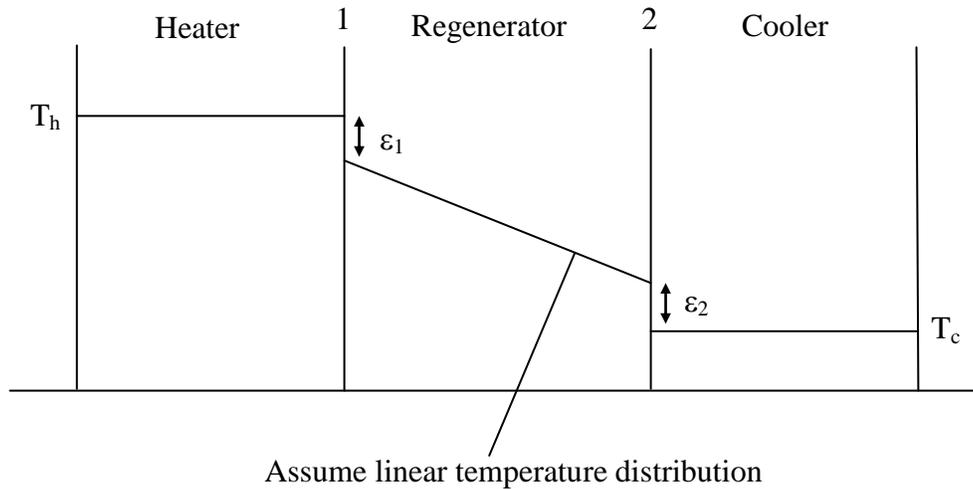
Regenerator constant eps2 (dimensionless)

0.0

Note: If you don't know any better you can use a value of 0.0 for eps1 and eps2. The values of eps1 and eps2 actually have little, if any, effect on the power output, and optimization. This means that they won't affect the optimal number of tubes and regenerator void volume resulting in maximum power. They almost exclusively affect just the heating and cooling energy required to produce a given power. However, even though they don't affect the "optimal" design it can still be insightful to change them to see the affect.

The effectiveness of the regenerator is given by ϵ_1 on the heater side, and ϵ_2 on the cooler side.

Typically, ϵ_1 (call this ϵ_1) and ϵ_2 (call this ϵ_2) are about 0.05-0.10. A graphical representation of ϵ_1 and ϵ_2 is given below.



ϵ_1 and ϵ_2 are convenient numbers used to account for non-ideal regenerator behaviour; in other words, “real” behaviour.

Looking at point 1 (from the figure above), the regenerator temperature is:

$$T_{\text{reg}} = (1 - \epsilon_1)T_h$$

Similarly, looking at point 2 the regenerator temperature is:

$$T_{\text{reg}} = (1 + \epsilon_2)T_c$$

The temperatures T_h , T_c and T_{reg} are in Kelvin.

As it turns out, small changes to ϵ_1 and ϵ_2 can result in large changes in heating and cooling requirements. However, only ϵ_1 is normally relevant in terms of cost, since the energy input is only with regards to the heater, which can use either a combustible fuel, or concentrated solar energy. The value ϵ_2 is normally less relevant since cooling is basically “free” given that we can use a radiator for cooling to the atmosphere. However, a low ϵ_2 value can still help minimize the size of the radiator and reduce cost on that end. This points towards the great importance of the regenerator for achieving high thermal efficiency (lower fuel cost) and minimizing the size and bulk of the cooling system.

In a nutshell, the penalty for non-zero values of ϵ_1 and ϵ_2 is more fuel for the heater and a bulkier cooling system.

For an ideal regenerator with 100% thermal efficiency, use $\varepsilon_1 = \varepsilon_2 = 0$. Looking at the graph, this means that the temperature of the regenerator matrix varies linearly from T_h at the heater end, to T_c at the cooler end.

Non-zero values of ε_1 and ε_2 are mainly the result of regenerator heat loss to the environment and axial conduction of heat along the regenerator matrix (in the flow direction).

(Note that my definition of ε_1 and ε_2 is opposite to the convention commonly found elsewhere; where an ideal regenerator is represented by $\varepsilon_1 = \varepsilon_2 = 1$).

Last input file parameter:

Scale factor to get the average pressure you want at the engine speed you want (default = 1.0)

1.0

This is a scale factor you enter after you have run the program once and looked at the (actual) average engine pressure in the results. As it turns out, the average pressure in the results is normally a bit different than the average pressure specified by the user in the input file. The scale factor is used to correct for this difference.

This discrepancy in average pressures is due to the difficulty in accurately calculating the mass of working gas in the engine so that the average engine pressure exactly equals what the user enters. The method used to estimate the mass of working gas is based on the *Schmidt Analysis* (not discussed here), which is a good approximation but does not result in an average pressure that is exactly equal to the specified average pressure. One way to deal with this discrepancy is to adjust the mass inside the program loop until the actual average pressure equals the specified average pressure. However, the problem with doing this is that the average pressure changes with different engine speeds. So, for which speed do you correct for the pressure difference? Only the user can answer this question.

This means that you (the user) must look at the results, and decide what engine speed best matches the result you are looking for. Your decision could be based on the speed at which there is maximum power, or whatever other criteria. Once you've decided, the next step is to match the average pressure at the speed you want with the average pressure in the input file. To do this, you have to manually calculate the scale factor. Fortunately, this is really easy to do.

For example, let's say you decide that an engine speed of 22 Hz is optimal. The actual average pressure at this speed is 8545000 Pa (from the output file). And the desired average pressure (from the input file) is 8900000 Pa. The scale factor is then: $8900000/8545000 = 1.0415$. Enter this value in the input file, save it, and then run the program again. After the program has run look at the results again. You will see that at 22

Hz (or a frequency close to it) the new average pressure will be very close to 8900000 Pa. It won't exactly match but it will be close enough.

Note that you can also use a scale factor that best matches the average pressures over a *range* of engine speeds. However, this works best if, within that range, the average pressures are close to each other.

Now, as an alternative to using the scale factor, some may think that it's better to enter an average pressure *higher* than the one you want, in the input file, knowing that the program results in an average pressure that's lower. Don't do this. The average pressure in the input file is used to calculate constants in the equations, so it must match the pressure you want. Otherwise you risk introducing excess error in the program calculations. If you don't fully understand this don't worry about it. Just keep in mind that correcting for the discrepancy between the average pressure in the input file and the results must only be done using the scale factor.

And lastly, when running the program for a new engine configuration, set the scale factor back to 1.0. Then, to correct for the discrepancy between average pressures, change the scale factor to the appropriate value (using the method explained above).

Using The Program (must-read)

The program file, called 'stirlingpro.exe' is easy to use. It is stand-alone and does not require any extra installations or change of settings on your computer to run. You simply put the program file into a directory of your choice, along with the input file, and then run it from the command line. For example, let's say the program and input file are located in a directory called 'MyStirling', and you want to run it from there. These are the steps:

(1) Go to the Windows Start menu. Select *Programs*, followed by *Accessories*, followed by *Command Prompt*. This will open the command prompt window.

(2) In the command prompt window set the directory to the 'MyStirling' directory (if you aren't already there). For example, if the directory is located at d:\John\MyStirling, then the first thing you do is type 'd:' at the command prompt (don't include the quotes). This will take you to the d: directory. Then at the prompt type: 'cd John\MyStirling'. This will take you to the right directory.

(3) Next, you run the program. To do this, type the following at the prompt: 'stirlingpro.exe', and the program will run. If you want to change the input parameters then edit the input file in Notepad and save it. Then run the program again. The new results will correspond to the new input parameters.

Alternatively, to run the program you can right-click 'stirlingpro.exe' and select Open. When the program runs it creates an output file called 'results'. This file contains the simulation results. These results can be plotted in a spreadsheet like Excel.

If (for whatever reason) the program “hangs” meaning it doesn’t finish running, type Ctrl-C (hold down Ctrl and press C). This will terminate the run.

Method of Solution (must-read)

Using the program is straightforward. You specify everything in the input file according to the details of your design, and then run the program. The results are calculated and output in a file called ‘results’. This file is created by the program.

The program can be used to optimize your design; meaning you can determine the optimal design that results in maximum engine power. The way to do this is to vary the number of tubes (in the heater and cooler) and regenerator void volume, in the input file, until maximum power is reached. Everything else in the input file is kept as constant.

To optimize your design for maximum power follow these steps:

- (1) Set the number of heater tubes to a low value, such as 10 or so.
- (2) Set the number of cooler tubes to a low value, such as 10 or so.
- (3) Set the void volume of the regenerator equal to $V_{\text{heater}} + V_{\text{cooler}}$ (described previously on page 29-30).
- (4) Set the scale factor to 1.0.
- (5) Set all the other input parameters to the values of your choice. Save the input file, and run the program.
- (6) Open the output file called ‘results’ (using Notepad or WordPad). Scroll down to the bottom. You will notice a chunk of data in which the power, thermal efficiency, torque, and average engine pressure are given at different frequencies (speeds). Make a note of the frequency at which maximum power occurs. At this frequency, make a note of the average engine pressure. Call this pressure P1. Calculate the following scale factor: (desired average pressure, from the input file)/P1. Enter this scale factor into the input file (leave everything else the same). Save the input file. Close the ‘results’ file.
- (7) Run the program again.
- (8) Once again, open the ‘results’ file. Look at the frequency at which maximum power occurs. The average pressure corresponding to this frequency should be very close to the desired average pressure in the input file. At this frequency, write down the maximum power, thermal efficiency, and average pressure for future reference. We are now ready to try the next configuration.
- (9) In the input file set the scale factor back to 1.0.

- (10) Increase the number of heater tubes by an increment of 10, or so.
- (11) Increase the number of cooler tubes by an increment of 10, or so.
- (12) Set the void volume of the regenerator so that it equals the new $V_{\text{heater}} + V_{\text{cooler}}$.
- (13) Save the input file. Close the ‘results’ file.
- (14) Run the program again.
- (15) Repeat steps (6) – (14) until the maximum power from step (8) starts to decrease. Once this happens you are done. You have gone past the point of maximum power, and the (optimal) configuration for maximum power is based on the previous run. In other words, the number of tubes, and regenerator void volume from the previous run will (approximately) be your optimal design values. These are what you will use for your design. If you want further clarity on this, Appendix G shows a specific example of the optimization procedure. It may be easier to follow that.

Note that thermal efficiency at maximum power stays roughly constant as you increase the number of tubes and regenerator void volume. So you don’t have to worry about optimizing thermal efficiency. This is explained in more detail later on.

In general, for each configuration you are running the program twice. The second run is to “match” up the average pressure (at the point of maximum power) with the average (desired) pressure specified in the input file.

Once you determine the optimal configuration you can now proceed to build the engine, keeping it as close as possible to the (optimized) design values. If there are any small changes during the building stage, you can see the effect of these changes by running the program again, using your “actual” design values as input.

The program takes less than a minute to run each time. So you’ll get your results quickly.

You may actually decide to use a smaller increment than 10 when increasing the number of tubes, especially when you know you’re getting close to the maximum power.

You might be wondering why I didn’t automate the optimization process so that you wouldn’t have to do it manually each time. I didn’t automate it because there may be different ways of adjusting heater, cooler, and regenerator void volume. For example, you might want to use different numbers of tubes for the heater and cooler. Or you might not want to make the regenerator void volume exactly equal to the volume of the heater plus cooler (based on their exposed lengths). It is not worth trying to accommodate all the different possibilities.

As a hint, try to keep the engine design as modular as possible, meaning that you can change and replace parts easily, as you work towards the best design. Afterwards, you

can choose to make a “permanent” engine that is less modular, once you are satisfied that you have made the best engine you can.

This next part is optional:

Prior to going through steps (1) - (15), it can be useful to first estimate the basic power your engine will produce given a certain swept volume and average pressure. Using the Beale formula (ref: http://en.wikipedia.org/wiki/Beale_number) one can roughly estimate the engine power given a certain swept volume and average engine pressure:

$$W = B_n PVf$$

Where:

W is the estimated power output (Watts)

B_n is the Beale number, which is usually in the range 0.11 - 0.15

P is the average engine pressure (Pa) over one cycle

V is the swept volume of the expansion space (m³)

f is the engine speed (Hz). A practical range of speed is, say, 10-30 Hz

The Beale formula is based on data observed from many different engines. In most instances these engines operate with a heater temperature of 650 degrees Celsius and a cooler temperature of 65 degrees Celsius (ref: “Stirling Engine Design Manual”, W. R. Martini, Second Edition. Page 99. The web link is: http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19830022057_1983022057.pdf).

From the Beale formula, one can estimate the power output given a certain swept volume and average engine pressure. However, since the Beale formula is a straight line when plotted against frequency, it is inadequate for higher engine speeds. Its main drawback is that it doesn't account for pumping losses, which become significant at higher engine speeds. In any case, it gives a good first estimate for power, based on only two design criteria. Simply adjust the swept volume and average pressure until you get the power you want, and then move on to steps (1) – (15). This allows you to set the basic power you want, and from there optimize the design.

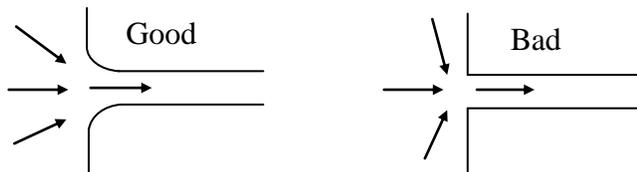
A Closer Look at the Heater and Cooler (must-read)

The following details are given in point form:

- You need a large number of tubes in the heater and cooler otherwise the engine won't start, or it will run at a low speed (due to high pumping losses). If your engine is running at a low speed, maybe a few hundred rpm, consider adding more heater and cooler tubes to speed it up, and increase power.
- The ratio 80:1 (ratio of exposed tube length to inside tube diameter) works well for all gases, according to my heat transfer calculations. But if you want to optimize for individual gases a closer look is necessary, since the ideal ratio may be a bit different depending on if you are using, say, air or hydrogen as the working gas. You will probably have to get your hands on some detailed empirical data in the literature, or do some experiments yourself in order to optimize the ratio, depending on the working gas you are using. However, if you are not using a regenerator, the ratio becomes much more dependent on the type of working gas used. For example, the absence of a regenerator would require a ratio of approximately 250:1 for hydrogen, and 325:1 for air. Hydrogen heats up faster than air so it requires less tube length.
- When using a regenerator, with an 80:1 ratio for the heater and cooler tubes, there is relatively little difference in the tube exit temperature for different gases. This is because the gas entry temperature is close to that of the heater and cooler walls (due to the pre-heating and pre-cooling of the gas by the regenerator). For example, with a regenerator, hydrogen may exit the heater tube at 650 degrees Celsius while air may exit at 630 degrees Celsius. But if there were no regenerator, the hydrogen may exit at 550 degrees Celsius while the air may exit at 480 degrees Celsius. The difference is much greater.
- You can do a sensitivity analysis on the ratio. For instance, change the ratio by a certain amount to see the effect on engine power and efficiency. For example, if you know that a ratio of 70:1 works better than 80:1, then use that for the program input and see what the effect is.
- The heater tubes/passageways are generally made of stainless steel because of its resistance to oxidation when exposed to high heat. The cooler tubes, which are usually cooled by water, can be made of aluminum since it won't rust.
- The heater and cooler tubes are usually thin-walled to allow heat to easily pass through.
- In general, the heater and cooler tubes must be as closely spaced together as possible (since there is usually a large number of them), but must also have adequate spacing between them to allow room for circulation. This is necessary

because the cooling medium (e.g. water) and heat from the heater must reach all the tubes, without being “blocked” by neighboring tubes.

- It is helpful to machine a small radius at the entrance of each tube/passageway in order to minimize “entrance losses” (which are a form of flow loss due to the gas rapidly changing direction as it flows into the tube/passageway). The figures below illustrate this.



- There are also “exit losses” which occur as the gas exits the tubes/passageways and rapidly slows down. These types of flow losses cannot be reduced using a small radius at the tube exit. Exit losses will always have the same magnitude, regardless if there’s a radius or sharp corner at the exit.
- Note that the inner (flow) surface of each tube/passageway is assumed “smooth” in the mathematics of the model. This also minimizes flow losses. During the fabrication process the inner surfaces will generally be made smooth anyway.

A Closer Look at the Regenerator (must-read)

The following details are given in point form:

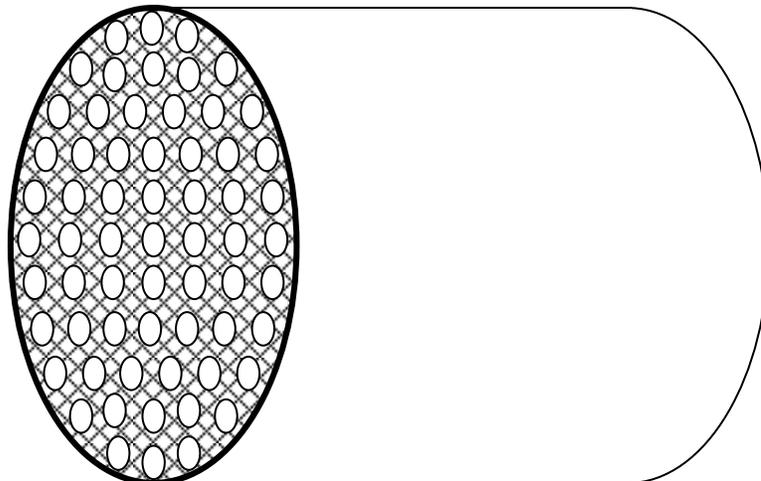
- The regenerator matrix material is usually stainless steel because of its oxidation resistance when exposed to high heat.
- Stainless steel wool is also an inexpensive regenerator material used by numerous Stirling engine enthusiasts.
- Metal screens are commonly used as matrix material. They are stacked in the (axial) flow direction, which works very well in limiting axial heat conduction, since heat is forced to travel *between* the screens. This helps to reduce losses in regenerator efficiency. Note that axial heat conduction in the regenerator is one of the loss mechanisms that result in non-zero values of ϵ_1 and ϵ_2 . In the literature, such losses are commonly referred to as “thermodynamic irreversibilities”.
- To further reduce axial heat conduction in the regenerator, it is desirable to use a matrix material with low thermal conductivity. Stainless steel is suitable in this regard.

- Thermal losses in the regenerator can also be reduced, by insulating the outside of the regenerator housing. This helps minimize heat loss to the environment.
- Two examples of good matrix material are: “400 mesh” which stands for “400 wires per inch”; and “Metex” formed from knitted steel wire. Both are made of stainless steel.
- In the model used, the regenerator pumping (flow) losses are assumed to be 80% of the *total* pumping losses caused by the flow of working gas through the heater, regenerator, and cooler. This is a reasonable estimate based on what I read in the literature [3]. This estimate generally corresponds to a matrix porosity of 65-80%, which is the porosity you should have in your regenerator (for best performance). The flow loss in the regenerator is always much higher than the flow loss in the heater and cooler tubes/passageways, simply because the working gas has to travel a more “difficult” path when flowing through it.
- As it turns out, the “optimal” engine design is not sensitive to the (assumed) percentage of pressure drop across the regenerator. So if 70% or 90% were used instead of 80%, the optimal design would not change. This is explained in greater detail later on, in the Sensitivity Analysis section.
- The purpose of the regenerator is to raise the thermal efficiency of the engine, and an effective regenerator (the design goal) will always contribute the most to pressure drop. Using this (constant) percentage of 80% for all designs is much easier than trying to explicitly model specific cases. Modeling the regenerator is much more difficult analytically than modeling the heater and cooler tubes. The 80% value is based on a review of optimal design data taken from the literature. Naturally, it is assumed that your goal is optimal design as well. However, at some point in the future I may decide to revisit the subject to see if a significant improvement can be made with regard to pressure drop calculation in the regenerator. That said; it is expected that the 80% assumption will still result in a good ballpark estimate of engine performance in general. To adequately model the regenerator one must obtain accurate information about it. In addition to porosity, one must know the size of the mesh holes, how the mesh screens are stacked relative to each other, their thickness, and their surface area. This is not nearly as straightforward as knowing the length and inside-diameter of a tube.
- In a nutshell, the 80% pressure drop assumption is based on the literature findings, which indicate 65-80% regenerator porosity (the main parameter), with a void volume roughly equal to the combined volume of the heater and cooler (based on their exposed lengths). In addition, the equations used in the literature to calculate flow loss in the heater and cooler tubes are basically the same as the equations I used. So there is a degree of consistency in my approach.
- The porosity of the matrix can be calculated as follows: First, calculate the volume that contains the matrix material (this is the volume of the regenerator

housing). Call this volume V_1 . For example, if the housing is in the shape of a cylinder, then V_1 is simply the volume of that cylinder. Next, measure the weight of the regenerator matrix. Next, divide this weight by the density of the matrix. This gives you the volume occupied by the matrix material. Call this volume V_2 . The porosity of the regenerator matrix is therefore: $(V_1 - V_2)/V_1$. Note that $(V_1 - V_2)$ is the regenerator void volume.

- It is worth mentioning that the proportion of pressure drop across the regenerator likely stays fairly constant over different engine speeds. In the following reference (page 12), the NASA model predicts that the regenerator pressure drop holds fairly constant at 70%, for speeds ranging from 30-60 Hz: “Comparative Analysis of a Stirling Heat Pump With Second-Order and Third-Order Computer Models”, N. Domingo, Oak Ridge National Laboratory, June 1985. The web link is: http://www.ornl.gov/sci/ees/etsd/btric/eere_research_reports/thermally_activated_technologies/engine_driven/stirling_rankine/modeling_and_simulation/ornl_con_162/ornl_con_162.pdf

By extension, if the proportion of pressure drop across the regenerator stays fairly constant for different engine speeds, this means that the proportion also stays fairly constant for different numbers of tubes, and (proportional) regenerator void volume. For example, doubling the engine speed is the same as decreasing the number of heater/cooler tubes and regenerator volume by half. In the latter case, the speed effectively doubles because the total flow area is half of what it was before. So from the point of view of the heater/cooler tubes and regenerator, there is no difference. However, there is an important assumption made here: It is assumed that the cross-sectional flow area of the regenerator is decreased by half as well, keeping its length the same. This is a reasonable assumption since, generally speaking, the regenerator flow area is what must increase or decrease in order to accommodate more or less tubes. Therefore, the regenerator length doesn't need to change. A good rule to follow is to make the cross-sectional flow area of the regenerator as small as possible while allowing enough room for the heater and cooler tubes to comfortably “fit inside”. The figure below illustrates this.



For some regenerator designs the heater and cooler tubes lie within a ring-shaped area. This means the regenerator must also be ring-shaped. The figure below shows a picture of the ring shaped regenerator used in the MOD II automotive engine (additional pictures of the engine are in Appendix E).



Reference: NASA Technical Reports Server
<http://ntrs.nasa.gov/search.jsp>

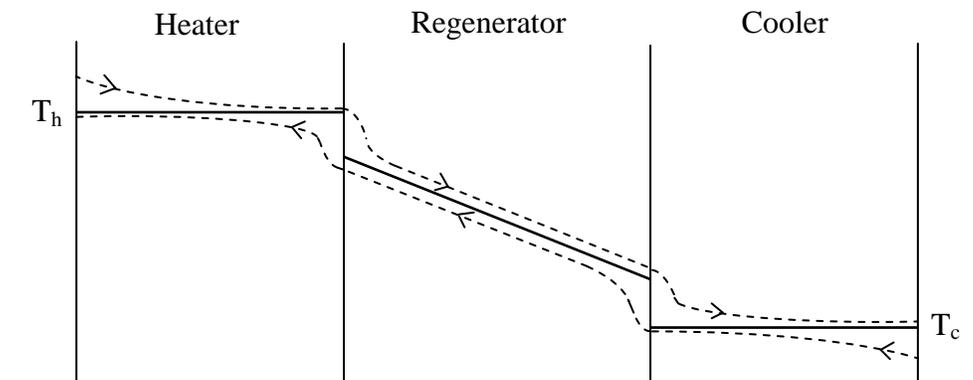
- Good heat exchange between the regenerator matrix and the working gas is a function of the size of the regenerator and the porosity. For instance, the working gas blows through a smaller regenerator *faster* and therefore needs to exchange *more* heat with the matrix per unit volume (in order to reach an exit temperature near to that of the heater and cooler). Conversely, the working gas blows through a larger regenerator *slower* and therefore needs to exchange *less* heat with the matrix per unit volume (in order to reach an exit temperature near to that of the heater and cooler). This means that the smaller regenerator requires a greater mass of matrix material per unit volume than does the larger regenerator, to have the same level of heat exchange with the working gas. So it follows that the smaller regenerator must have *lower* porosity than the larger regenerator. But it doesn't end there. To ensure a good design, one must also account for the effect of pumping losses and dead volume. This is discussed in the next point.
- The volume and porosity of the regenerator must be chosen to maintain the delicate balance between requiring too much pumping power and adding too

much dead volume, while providing good heat exchange properties. For example, a dense matrix (low porosity) in a smaller regenerator volume would provide good heat exchange with the working gas, but may require too much pumping power. But a less dense matrix (higher porosity) in a larger regenerator volume would also provide good heat exchange with the working gas, but it may add too much dead volume. So there is an optimal middle ground. However, since matrix materials are usually metallic, they have a very high volumetric heat capacity relative to the working gas. So as long as the regenerator porosity falls within a certain range, it will function well in general. Looking at the literature, a good range of porosity of the regenerator for optimal performance is 65-80%. This means that the matrix material fills 20-35% of the regenerator volume.

- From the study described in reference [5], an ideal regenerator design consists of a matrix porosity of about 80% and a matrix wire diameter of about 50 μm (50×10^{-6} m).
- As mentioned, a regenerator is required in order to achieve high heater and cooler efficiency. In fact, a 1% improvement in regenerator efficiency can improve the overall engine efficiency by several percent. This points to the sensitivity of engine efficiency on regenerator performance.
- To gain a feel for the effect of regenerator size on power and efficiency, you can do a sensitivity analysis. Make changes in the regenerator void volume and observe the effect on engine power and thermal efficiency.

Graph to help you visualize the temperature of the working gas as it flows through the heater, regenerator, and cooler (must-read)

The figure below may help you to better visualize the temperature of the working gas as it flows in both directions through the heater, regenerator, and cooler. The solid line represents temperature. The dotted line represents the working gas.



Engine Losses (must-read)

In addition to thermal losses due to regenerator inefficiency (discussed previously) there are also other engine losses to consider. The details of these are given in point form below:

- In a real engine there are friction losses, such as in the mechanical drive, linkages, and between the piston/displacer seals and cylinder wall. This directly reduces engine power. These friction losses can only be accurately calculated with experimental measurements. They are not accounted for in the program.
- There are also thermodynamic losses such as from hysteresis effects, due to compression of the working gas in the expansion and compression space, causing it to heat up to a temperature higher than that of the cylinder wall, during parts of the cycle. As a result, heat is lost to the environment. This loss mechanism can be minimized, by insulating the outside walls of the expansion and compression space. In the program it is assumed that the expansion and compression space are adiabatic, which means that the working gas does not lose heat through the cylinder walls. This is a good assumption for large high-pressure engines.
- Other losses include: working gas leaking out of the engine, heat transfer inefficiency from heat source to heater tubes, and other thermodynamic inefficiencies due to heat loss in other parts of the engine. For instance, there are heat transfer losses that occur as a result of heat flowing along the engine wall from the hot side to the cold side. There are heating losses that occur between the gap of the displacer and the cylinder wall, due to the temperature difference between the expansion and compression space. No provision is made in the program to account for these losses. For the most part they can only be accurately calculated by experimental measurements, and then minimized by proper material selection and design.
- As mentioned, regenerator inefficiency is one of the major sources of thermal loss. But the other thermal losses mentioned above can (in combination) further reduce thermal efficiency by several percent.
- This program is meant to optimize the design based on the *intrinsic* engine thermodynamics, which models the main physical phenomenon occurring *inside* the engine. For the most part, the losses mentioned in the previous paragraph affect the thermal efficiency only (i.e. by reducing it). This means that extra heat energy input is required to compensate for these losses. In other words, the engine power itself is not affected, provided there is sufficient heat energy available to compensate for the thermal losses.
- It is very important to know that (with the exception of hysteresis losses and leakage of working gas), accounting for all the above-mentioned losses would *not* affect the thermodynamics and physics *inside* the engine. So for optimization

purposes they can be excluded from the model. In other words, their exclusion will not affect the number of tubes and (proportional) regenerator volume required for maximum power.

- One way to significantly improve thermal efficiency in the design is to improve the heat transfer efficiency from heat source to heater tubes. A common loss mechanism in this regard is heat loss to the surrounding environment (e.g. warm exhaust from a burner). A way to minimize this loss is with an air Preheater. Using the exhaust stream, a Preheater heats the air before it enters the combustion chamber, and more of the heat energy of the fuel is used. This is also more economical since it reduces fuel consumption. In addition, you can also minimize heat loss by placing an insulated enclosure around the heat source.
- A well-designed heat source, such as burner with air Preheater, can have a heat transfer efficiency of 90%. This means that 10% of the heat is lost to the environment. This loss further reduces thermal efficiency by several percent. For example, an engine operating at 40% thermal efficiency with (theoretically) perfect heat transfer from the heat source, would run at 36% efficiency with 90% heat transfer efficiency (0.90×0.40).
- One interesting loss mechanism not mentioned earlier, is *Adiabatic Loss* [1]. This loss results from a high compression ratio – (maximum-engine-volume)/(minimum-engine-volume), which forces the gas temperature in the expansion and compression space (during parts of the cycle) to exceed the heater and cooler temperature, respectively. This results in heat being pumped *out* of the heater and cooler due to the positive temperature difference. This lowers thermal efficiency. But at the same time a sufficiently high compression ratio is necessary for high power. However, if it's too high, the adiabatic loss becomes excessive. This is another example of the conflicting requirements in Stirling engine design.

Note that the Adiabatic Loss is inherently captured by the equations of the program, and doesn't need to be accounted for explicitly.

A Closer Look at Power and Efficiency (must-read)

The following details are given in point form:

- It may be possible that with no regenerator (and longer heater/cooler tubes to compensate), you will be able to achieve a higher power output due to lower pumping (flow) losses. But, your thermal efficiency will be very poor as a result, and you will need much more fuel (heat input) to produce power. This will overwhelm any power advantage gained.
- It's better to use a regenerator to achieve high thermal efficiency even at the expense of lower power. You can always scale up the engine size to get additional

power, and maintain the high thermal efficiency. You end up using much less fuel that way. For example, without a regenerator you may have twice as much power than with a regenerator, but your thermal efficiency may be three times less. So it's much better to just scale up the size of the engine or (in some cases) just increase the engine pressure to get the additional power, without sacrificing efficiency. However, in some rare cases keeping engine bulk to a minimum is very important. So to achieve that, some compromise will have to be made at the expense of efficiency. For example, it's the same idea as using a two-stroke gasoline engine for applications demanding a high power-to-weight ratio. A two-stroke engine has twice the power of a four-stroke engine but is consequently much more polluting and has lower fuel efficiency.

- If the available heat energy is very high relative to the power output that means the engine pressure should be increased in order to reach its full power potential. As a result, thermal efficiency will also increase. For example, if the available heat energy from the heat source is 10,000 W and the power output of the engine is 25 W, we can definitely get more power by increasing engine pressure. Increasing pressure enables the engine to use more of the available heat energy.
- Let's examine the previous point more closely. Increasing the engine pressure will result in a speed and power increase. But it also results in two other direct consequences. First, the heater wall temperature will *decrease* since heat energy is being "pumped" away at a faster rate (due to the higher density of the working gas as it passes through the heater). Second, the cooler wall temperature will *increase* because the cooler has a higher cooling demand placed on it (due to the higher density of the working gas as it passes through the cooler). And if the engine pressure continues to increase, the heater temperature will continue to decrease and the cooler temperature will continue to increase. This will eventually result in a sharp drop in thermal efficiency, and as a result the power and speed of the engine will drop as well. This is a result of Carnot efficiency (discussed in Appendix D). One can compensate for this by using a control system to pump more fuel into the heater to bring the heater temperature back up to the target temperature (also known as the set point temperature). Furthermore, it will also be necessary to pump the coolant through the cooler at a faster rate since heat must be removed faster to maintain the target (set point) temperature of the cooler.
- However, if the source of heat energy is constant and cannot be regulated, such as with a wood firebox, or concentrated solar energy, then the size of the engine and its operating pressure must be chosen to best match the (constant) heat input. The easiest way to build an engine, given this condition, is to first choose the engine power you want given the maximum permissible heater temperature, based on design limits (maximizing heater temperature is desirable because it results in maximum thermal efficiency). You then calculate the heat input requirement (e.g. in Watts), and size the heater accordingly. For example, this may involve choosing the correct size of solar concentrator (e.g. parabolic mirror), knowing

that incident solar radiation is approximately 1000 W/m^2 , and knowing that a certain size of mirror is necessary to produce the required temperature (at the focal point).

- The compression ratio – (maximum-engine-volume)/(minimum-engine-volume) – for maximum power and thermal efficiency, is usually less than 2. Note that the compression ratio is a secondary result of your engine dimensions, which is based on swept volume and dead volume. It is not specified explicitly in the program.
- One interesting point is that, by lowering the heater temperature (decreasing the temperature difference between the hot and cold source), the number of heater and cooler tubes required for maximum power increases. This means that the compression ratio becomes lower as a result. Conversely, higher heater temperature (greater temperature difference) requires a *lower* number of tubes for maximum power. And the compression ratio becomes higher as a result. This is an interesting consequence of the physics of Stirling engines.
- Note that some sources refer to compression ratio as the ratio of swept volume of the displacer to the swept volume of the power piston (in beta and gamma engines). Other sources refer to the compression ratio as (maximum-volume)/(minimum-volume) in either the expansion or compression space. This is a common convention for Internal Combustion engines.
- For engines with very high pressure it takes a lot of torque to get them started because you have to move them past the compression stage. A way to get around this is to lower the engine pressure at the start. And then once the engine starts running increase the pressure. In some cases it may be best to use an electric starter with high torque.
- In my DIY Stirling engine project (see [Stirling_engine_project.pdf](#)), I used only one tube for the heater and cooler. This was not optimal. The speed could never reach a high level because it would take too much pumping power to push the air through a single tube. It would have been better to use many narrow diameter tubes. But in doing this, you move out of the DIY realm and into the realm of needing a machine shop and specialized materials. This program was developed afterwards, so it does help to have the benefit of hindsight.
- The rotational speed of a Stirling engine at zero shaft-load is a function of engine losses caused by mechanical friction in the linkages, bearings, seals, etc., and thermodynamic losses *inside* the engine, such as pumping (flow) losses. The engine will accelerate until "equilibrium" is reached (constant rotational speed). This is the point at which the engine power is balanced by the losses. In other words, the following equality applies in the case of constant engine speed: **mechanical and thermodynamic losses = engine power**. In the case where you have a load applied to the engine (such as a generator) the engine will reach a speed *lower* than the zero shaft-load speed. This corresponds to the same basic

equality, but with an extra term added: **mechanical and thermodynamic losses + generator load = engine power**. As mentioned, the engine speed at which this equality holds is the *final* rotational speed of the engine. Note that some of the variables on the left side of the equality, such as pumping losses, are strongly related to engine speed. But other variables, such as mechanical friction and generator load, may not be as strongly related. Ideally, you want this equality to correspond to high generator output while maintaining good thermal efficiency, because that means you get the most “bang for your buck”.

- It is a good idea to keep the mass of the pistons, displacers and linkages as strong and lightweight as possible in order to not burden the engine in accelerating these components, during parts of the cycle. To make the pistons and displacers as strong and lightweight as possible they must be hollowed out on the inside, while leaving enough material on the perimeter to provide adequate strength. However, flywheel design requires different design criteria. Flywheel mass must be substantial in order to move the engine past the compression part of the cycle. A hefty flywheel keeps the engine running “smooth”.
- The phase angle for maximum power and thermal efficiency occurs at roughly 90°. It is possible to increase the power slightly by using a phase angle of around 80°, but then your thermal efficiency decreases slightly. Therefore, the optimal phase angle for all-around performance is 90°.
- High-power engines typically have heater temperatures in the range of 650 - 800 degrees Celsius. This is necessary for high power density and thermal efficiency.
- If possible, it may be a good idea to have an adjustable swept volume, or phase angle in your design, as this will allow you to control engine power, perhaps more easily than by adjusting engine pressure, or heater/cooler temperature. An easy way to adjust swept volume is by varying the stroke length in the piston or displacer.

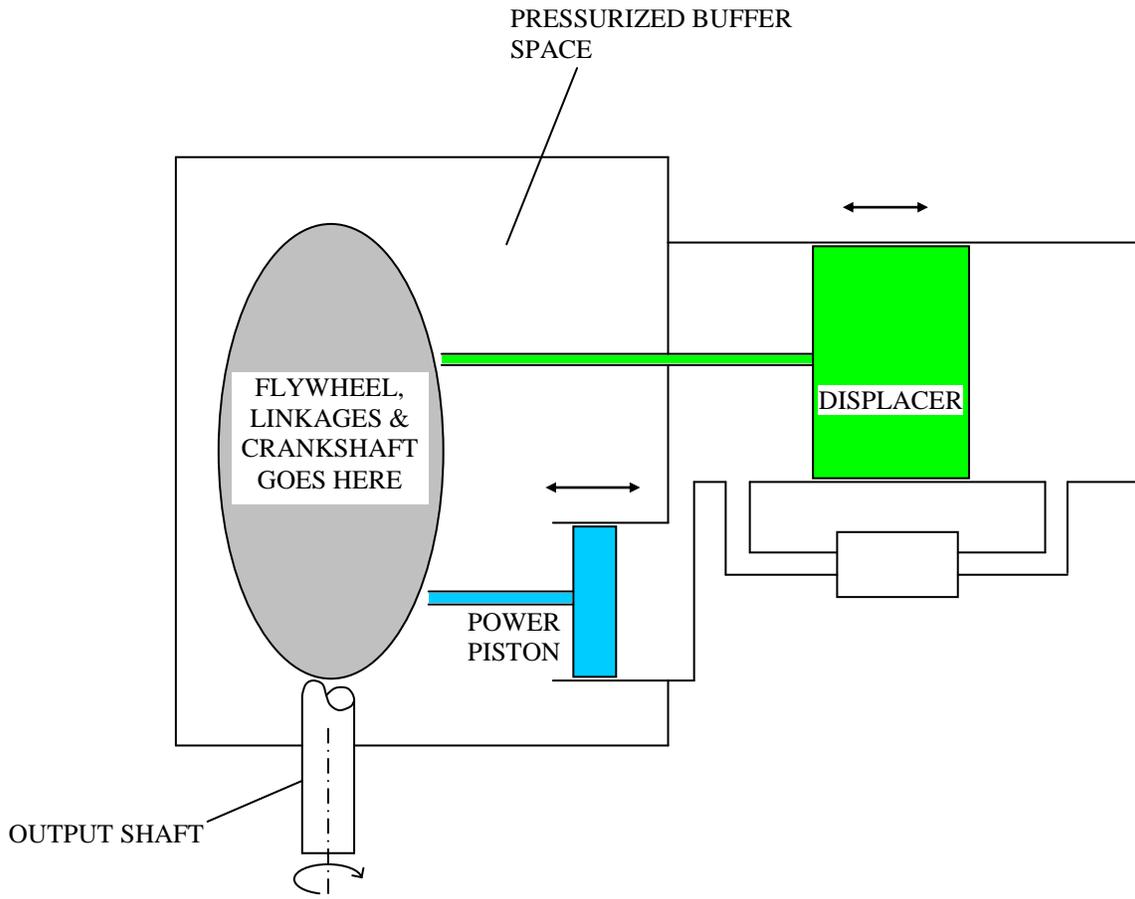
Buffer Space (must-read)

- There are designs that use a sealed buffer space (also called a “bounce” space) containing the flywheel and crankshaft/linkage mechanism. The buffer space is located *external* to the engine and creates a boundary between the atmosphere and the regions where gas can leak (at the interface between piston seals and cylinder wall). The buffer space is normally pressurized to the average engine pressure. The buffer space is large compared to the swept volume of the pistons in order to minimize pressure fluctuations inside of it. A pressurized buffer space reduces the requirements of the pistons and displacer rod seals, which only have to withstand pressure fluctuations above and below the buffer space pressure. These pressure fluctuations are much smaller than the pressure difference that would result if

there were no buffer space, and the pistons were directly exposed to the atmosphere.

- A pressurized buffer space reduces the strength requirements of the crankshaft and linkages because of the smaller pressure difference between the inside and outside of the engine. In addition the flywheel mass can be much less, for the same reason. This is explained in more detail in the Flywheel Design section.
- A pressurized buffer space also dramatically reduces the starting torque needed to get the engine running.
- Note that the power of the engine does not change with the use of a pressurized buffer space. The mathematical proof of this is given on page 116.
- A proper lubricant must be used that can withstand the high pressure of the buffer space. And care must be taken that it doesn't leak into the engine. This can happen during the parts of the cycle where the engine pressure is less than the buffer space pressure.
- In the cases where the engine drives a generator, it can be placed inside the buffer space. The buffer space can then be completely sealed using static seals, including the hole where the power cable comes out. However, in the cases where the output shaft protrudes from the buffer space, the gap between the shaft and through-hole must be sealed with dynamic seals designed for rotating components. Such seals tend to resist pressure better than the dynamic sliding seals used for pistons and displacers.

An illustration of the buffer space is given below, using a gamma engine as an example.



A Closer Look At The Displacer (must-read)

The displacer is generally made of stainless steel because of its resistance to oxidation when exposed to high heat.

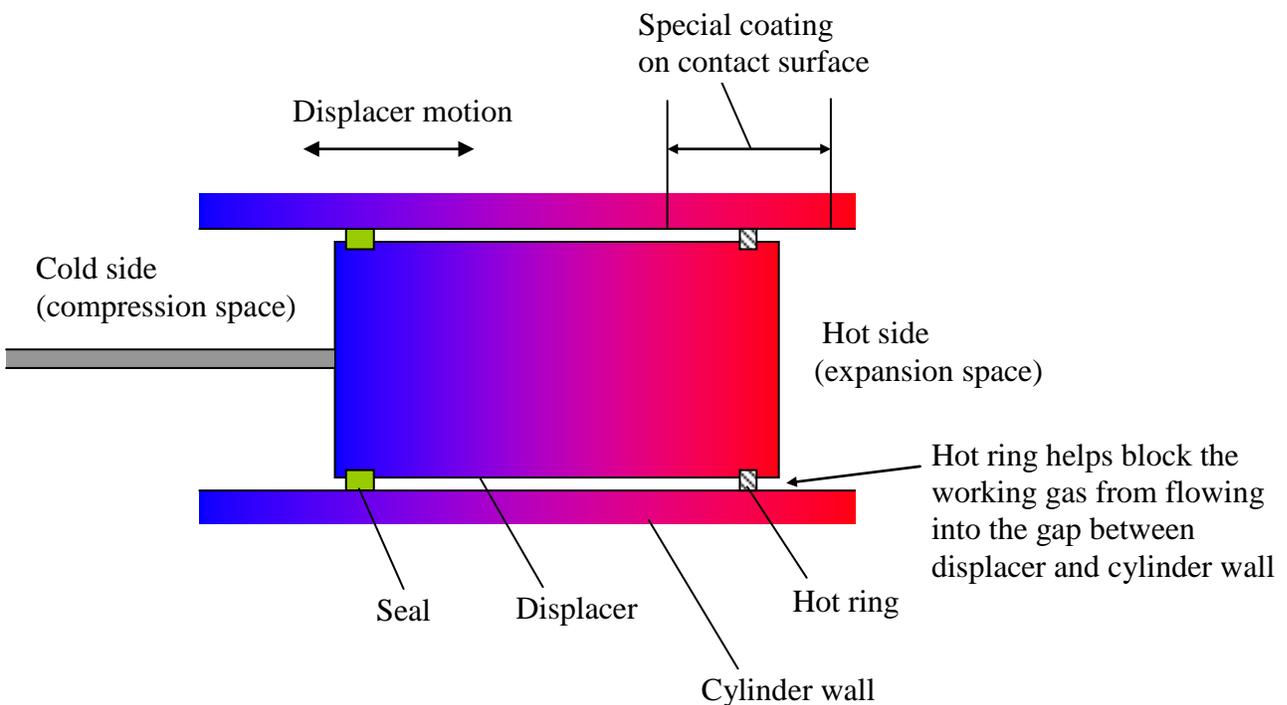
For beta and gamma engines, the main purpose of the displacer is to force the working gas through the heater-regenerator-cooler assembly. And depending on the direction of travel of the displacer, the working gas will exit either cool or hot. The result of this is power production in much the same way as an alpha engine.

In beta and gamma engines, the annular gap between the displacer and cylinder wall is typically sealed, using seals on the cold end (compression space side). This helps reduce pumping (flow) losses resulting from the working gas “squeezing” through the gap.

The difference in diameter between cylinder ID and displacer OD is typically less than 1 mm. The annular gap must be small to minimize dead volume. In addition the gap must also be long enough to allow the hot working gas from the expansion space to cool sufficiently by the time it reaches the seal (located on the compression space side). This is important because seals generally cannot withstand high temperature. (In the MOD II engine (Appendix E) the displacer length is about 1.5 times the diameter – this provides a sufficiently long gap).

Unfortunately, the penalty for having a long gap is a loss in thermal efficiency of the engine. This is due to the heat loss that occurs when hot gas flows into the gap and loses heat to the (colder) displacer and cylinder wall. As the gas flows into the gap it touches progressively “colder” surfaces on the displacer and cylinder wall. And when the gas flows back out of the gap it is colder than before, and heat energy is lost in the process. In essence this is a thermal “short-circuit”. Heat energy which would otherwise be used to produce power, is lost due to the round-trip motion of the gas into and out of the gap. Therefore, additional heat energy must be provided by the heater to compensate.

However, it is possible to minimize this loss by placing a metal ring (called a “hot ring”) around the displacer at the hot end. This helps prevent hot working gas from flowing into the gap during the displacer motion. Such a ring must be able to operate well without lubrication, at high temperatures. Perhaps a wear resistant alloy that functions well at high temperature can be used for the ring material. In addition, a self-lubricating coating can be applied on the cylinder wall, within the contact region. The figure below shows the basic set up. The temperature gradient is shown visually as blue (cold) transitioning to red (hot).



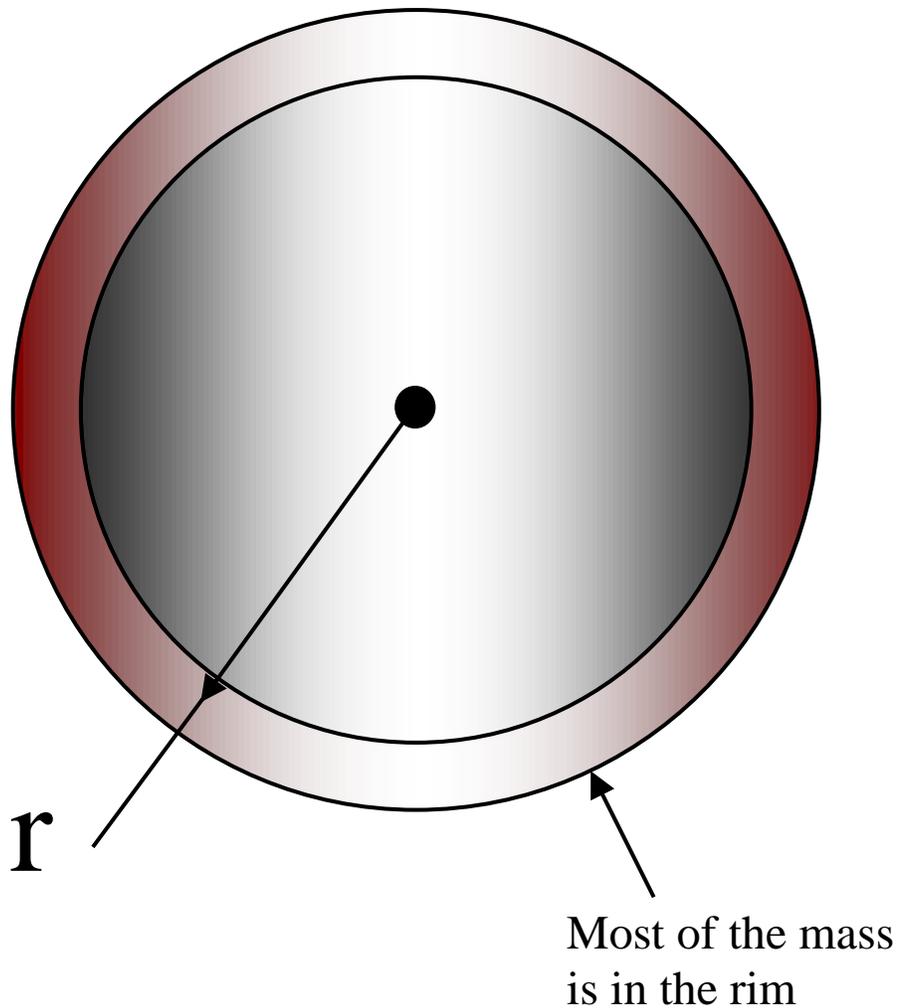
In a study done by NASA in 1988, tests were performed using a cobalt-based alloy (Stellite 6B) as the hot ring material, and PS200 as the coating (PS200 is a metal-bonded chromium carbide matrix with dispersed solid lubricants). It was found that fuel consumption in the heater was reduced by 3-7%. This study can be found at: http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19880006490_1988006490.pdf.

It is interesting to know that the above configuration for a displacer can also be used for the expansion space piston in an alpha engine. By using a long piston and a hot ring, the piston seals can be protected from the hot gas in the same way that displacer seals are protected. But it is necessary for the expansion space piston to be properly ventilated on the outside.

Flywheel Design (must-read)

Selecting a flywheel of suitable size and mass is an important aspect of the design. It has to be “heavy” enough to absorb energy from the engine during the power stroke (without speeding up too much) and then use that energy to “push” through the compression stroke (without slowing down too much). A proper flywheel will minimize fluctuations in rotational speed. This is necessary to ensure “smooth” engine operation. If the flywheel is not heavy enough the engine rotation will be “jerky” and engine performance will be poor.

The goal is to design a flywheel that is heavy enough to do the job but is not too big and heavy. When I speak of “heavy” I am specifically referring to the rotational-inertia of the flywheel about its center of rotation. This is a specific physical quantity related to the mass and size of the flywheel. We want a flywheel that has high rotational-inertia while keeping the mass as low as possible. This is best accomplished by using a flywheel with a solid outer rim, in which most of the mass is concentrated around the outside. The picture below shows a picture of such a flywheel.



The radius of the flywheel at the mid-point of the rim is given by “ r ”.

Mathematically, the rotational-inertia of the flywheel is given by:

$$I = mr^2 \quad (*)$$

where m is the mass of the outer rim. Note that this equation is marked with (*). It will be referred back to later on.

This equation is straightforward to use. If you know what the rotational-inertia (I) has to be for your particular engine design, you can choose a value of m and r such that the above equation is satisfied.

A good rule of thumb is to put 90% of the total flywheel mass in the outer rim. The remaining 10% of the mass is in the inside hub, for support. It can be ignored when using

the above equation. Ignoring this remaining 10% means that your *actual* rotational-inertia will be slightly higher than what you designed for.

You can estimate what I should be, based on the compression work required to “push” the engine from the position of maximum-internal-volume to minimum-internal-volume. Note that “internal volume” is defined as the total volume inside the engine.

Before getting into the details for calculating I , let’s look more closely at my definition of “compression work”. You can skip this next part if you like.

A Quick Aside (optional)

It is important to know that “compression work” as stated above, does not refer to the compression work in the compression space. It is a different quantity. When calculating the compression work necessary to “push” the engine from maximum-internal-volume to minimum-internal-volume, it is necessary to consider the combined effect of the compression space and expansion space. It follows that, mathematically speaking, the following sum must be calculated as the engine moves from maximum to minimum volume: (expansion space change in energy) + (compression space change in energy). The sum of these two is the compression work. To help visualize this, imagine that the engine is at maximum internal volume, and you turn the crankshaft by hand until minimum internal volume is reached. Assuming that the engine pressure is always above atmospheric throughout the rotation, you will feel a resistance. This resistance is (unavoidably) a direct result of the combined change in volume of the compression and expansion space (due to the piston/displacer motions).

Therefore, the compression work required to move the engine from maximum volume to minimum volume is the total energy that you used to turn the crankshaft. This is the energy required to physically “compress” the working gas inside the engine, plus overcome any load resistance (such as from a generator connected to the output shaft). This energy is provided by the flywheel as it rotates.

Now, as the engine moves from minimum-internal-volume to maximum-internal-volume energy is output from the engine. Let’s call this the “expansion work”. This is equal to the energy gained by the flywheel. If there is no load, subtracting compression work from the expansion work gives positive net energy. However, if there is a load, and the engine (and flywheel) is rotating at constant speed, the expansion work will equal the compression work. This means that the flywheel loses and gains the same amount of energy in the compression and expansion stages.

For positive net energy output the flywheel will be at a higher rotational speed at the end of the expansion cycle. Therefore, the engine will accelerate and turn faster until a final equilibrium speed is reached (depending on the loading and pumping losses inside the engine).

The angle of rotation of the crankshaft, to move from maximum-volume to minimum-volume, is 180° . And to move from minimum-volume to maximum-volume the crankshaft must rotate 180° to 360° (this takes it back to the starting position). In other words, it takes half a crankshaft rotation to go from one extreme to the other.

The above is an intuitive explanation. It assumes that the engine pressure is always above atmospheric pressure throughout the rotation. If this is the case, the compression work will always take place from 0 to 180° . And the expansion work will always take place from 180° to 360° .

However, if you are using a buffer space with pressure equal to the average engine pressure, the above description no longer applies. The net work produced by the engine will still be the same, but the expansion and compression stages will follow different profiles. This is because the pressure inside the engine goes above and below the buffer space pressure, in stages. This complicates the trends. As a result, compression work no longer takes place strictly between 0 and 180° , and expansion work no longer takes place strictly between 180° and 360° . Instead, compression work might take place between 120° - 180° , and 270° - 360° , while expansion work takes place over the remaining intervals. Fortunately, we don't need to worry about the details of this. It is only important to know that the net work at the end of a cycle ends up being exactly the same.

End of Aside

You can calculate the value of I based on several criteria that I think is reasonable for high power engines. These criteria are:

- The flywheel speed is allowed to vary ± 1 Hz at an average engine speed of 20 Hz. For slower engine speeds this variation will be higher, and at faster engine speeds this variation will be less. In other words, the engine runs “smoother” the faster it rotates.
- The calculation applies to alpha, beta, and gamma configurations as shown in this manual. The calculation for I does not apply to multi-cylinder configurations.
- The engine is spinning at constant speed, meaning there is a load, such as a generator, attached to the output shaft.

The value of I is given below, and is based on two criteria: (1) A pressurized buffer space is used, with pressure equal to the average engine pressure, and (2) There is no pressurized buffer space; meaning that the outside faces of the pistons are exposed to atmospheric pressure.

Pressurized Buffer Space at Average Engine Pressure

$$I = \frac{W_c}{2400} \quad (\text{units in kg}\cdot\text{m}^2)$$

No Pressurized Buffer Space

$$I = \frac{W_c}{800} \quad (\text{units in kg}\cdot\text{m}^2)$$

where W_c is equal to the compression work plus half the net work per cycle, given in the 'results' file.

The above values for I are calculated based on: 1) the permissible speed threshold of ± 1 Hz at an average engine speed of 20 Hz, and 2) the equation of energy for a rotating object (which is a flywheel in this case), and knowing that the change in energy must equal W_c . The derivation of this is not given here.

Note that the value of I corresponding to a pressurized buffer space is much lower than the value of I corresponding to no buffer space. This is because, with a pressurized buffer space, the forces acting on the flywheel are much lower (since the difference in pressure between the inside and outside of the engine is less). This results in lower compression work. As a result, the flywheel doesn't have to be as "heavy" to maintain a certain speed threshold. This illustrates a clear advantage of using a pressurized buffer space.

Sample Calculation

Let's assume we are using a pressurized buffer space.

From the 'results' file, the values for compression work and net work can be chosen at any frequency (the reason for this will be given below). For example, from the 'results' file (at 15 Hz) locate the following two items:

"The compression work for $V_{\text{max-to-}V_{\text{min}}}$ (J) = -1250.237415972".

and

"The net work per cycle (J) = 1600.297327511"

The compression work is given as a negative number, so we must convert it to a positive number → 1250.

Therefore, $W_c = 1250 + 1600/2 = 2050 \text{ J}$.

As it turns out, the compression work plus half the net work is roughly constant, no matter what engine frequency you choose these values from. For example, at an engine frequency of 7 Hz, $W_c = 2050 \text{ J}$ (the same). This is a very useful fact! Another very useful fact is that W_c is the same even if the assumed percentage of pressure drop across the regenerator is changed to, say, 70% or 90%. Therefore, you don't have to worry about how sensitive W_c is to the regenerator pressure drop. It isn't!

From the above equation for a pressurized buffer space we calculate:

$$I = 0.854 \text{ kg}\cdot\text{m}^2$$

Next, let's assume we want a flywheel radius of 20 cm. This must be converted to meters, so $r = 0.2 \text{ m}$.

Substitute I into equation (*) and solve for m :

$$0.854 = m \times (0.2)^2$$

Therefore, $m = 21.4 \text{ kg}$. This is the mass on the rim of the flywheel. Given that there is mass on the inside hub as well, the actual value of I will be higher. This means that the engine will run even "smoother" than predicted.

This mass of 21.4 kg might seem a bit high, but for $W_c = 2050 \text{ J}$, the engine will have at least a few horsepower.

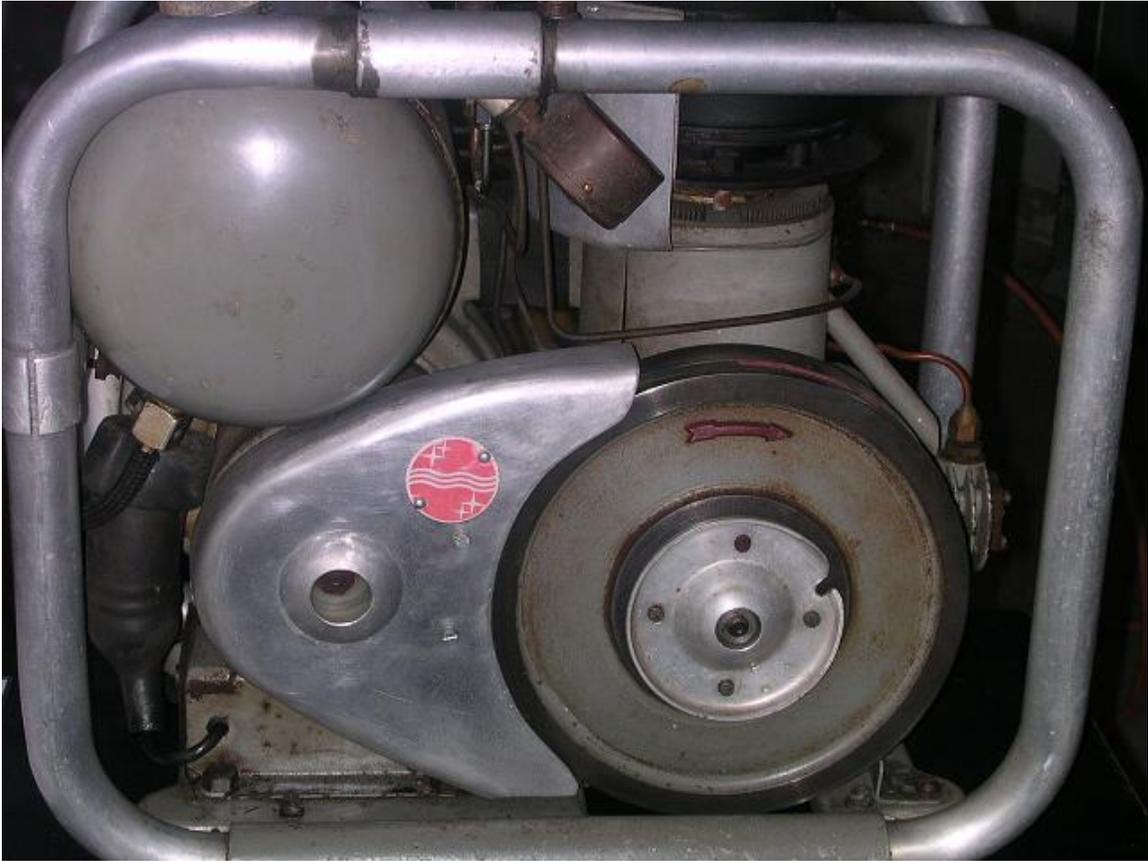
It is possible to further reduce the mass of the flywheel using a larger flywheel radius r . However, the radius must only be as large as practical.

Another way to further reduce the mass of the flywheel is to use multi-cylinder arrangements. Automobile engines, for example, can use lighter flywheels than single stroke engines. The reason is because they use multiple pistons and cylinders, connected to the same crankshaft (offset by an optimal amount). When one piston is in the "compression" part of its stroke another piston is assisting it during the "expansion" part of its stroke. Therefore, the flywheel doesn't have to store as much energy to smoothly move the engine past the compression stages, and it can be made lighter as a result.

The picture below shows a 200 W Philips Stirling engine generator, produced in 1951.

Notice that the flywheel has most of the mass on the rim (source:

http://en.wikipedia.org/wiki/Stirling_engine).

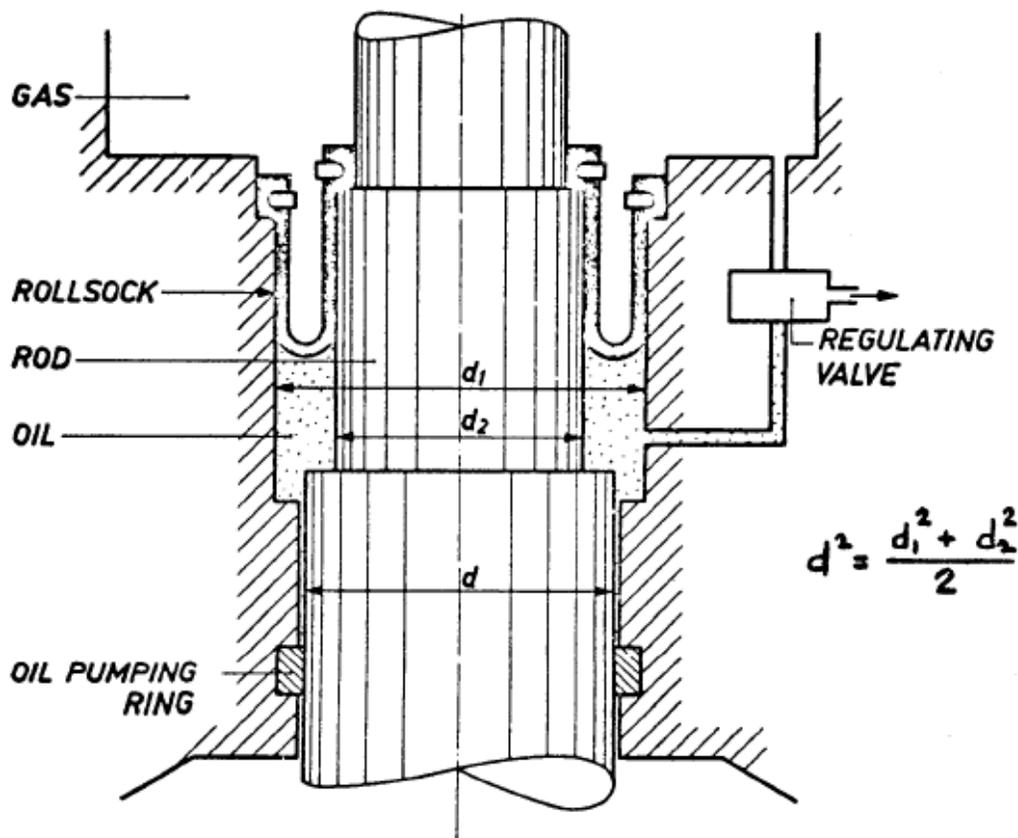


Miscellaneous Items (must-read)

- The engine components that are exposed to high heat are generally made of stainless steel. These include: 1) expansion space piston (in alpha engines), 2) expansion space cylinder, 3) displacer (in beta and gamma engines), 4) heater, 5) regenerator matrix, 6) regenerator housing.
- The best working gas to use for achieving high power and thermal efficiency is hydrogen. Helium is the second best. However, helium is probably the safer choice, especially for those not wishing to take “explosive” risks.
- The program is best suited for modeling reasonably sized power producing engines, not the small hobby size engines you typically see on YouTube. The main reason for this is as follows: Small hobby size engines don’t generally use tubes/passageways for the heater and cooler. The heating and cooling takes place at the surface area of the expansion and compression space. In the program, the expansion and compression space is treated as adiabatic, meaning there is no heat transfer with the environment. This is generally a good approximation for engines with large expansion and compression space volumes relative to their surface

area. But in small hobby engines this assumption doesn't hold true because the surface area of the expansion and compression space is large relative to their volume. For this reason these surfaces can be used as effective heat exchangers, by directly exposing them to a hot and cold source.

- The Ideal Stirling cycle is very different from the cycle for practical (real) engines. In many descriptions of Stirling engines a single working space is given, and the thermodynamic analysis follows from that. The working gas is expanded and compressed using a single piston. And the heating and cooling takes place at the outside surface. Although it works well as a basic description of the Stirling cycle, this analysis cannot be directly applied to real power-producing engines. In a real engine you need a separate expansion and compression space, in which the working gas is shuttled back and forth through heat exchangers. This is the only way to enable the rapid heating and cooling of the working gas necessary for high-power engines, operating at high speeds. Therefore, the Ideal cycle analysis cannot be used to model real engines.
- The real engine will always have some leakage, however small. If using air as the working gas you can have a small compressor to maintain engine pressure. If using a gas such as hydrogen or helium it's best to use a high-pressure storage tank to maintain pressure. Once the engine pressure drops below a certain value the tank releases gas into the engine until pressure is restored. If hydrogen is used one can even have an electrolysis device, which can produce hydrogen from water using electricity. This makes it easy to replace gas that has leaked.
- "Pumping Leningrader" sliding seals (in the MOD II engine) and Roll sock seals have been used with great success, lasting for thousands of hours during operation. PTFE (Teflon) seals can also be used, provided they are not exposed to very high temperature. The picture below shows a diagram of a Roll sock seal. This picture is taken from page 21 of: "Historical Review of Stirling Engine Development in the United States from 1960 to 1970", Energy Research and Development Agency, W. H. Percival, July 1974 (ref: http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19780072756_1978072756.pdf). This report also provides detailed information on the different seals used in Stirling engines, and how well they perform.



- When using oil for lubrication, make sure that it doesn't seep into the engine. The heat and pressure inside the engine will evaporate the oil and quickly foul the heat exchangers as a result. To combat this, some pistons use a combination of different seals, both to prevent gas from leaking out of the engine and to prevent oil from seeping in.
- For a beta and gamma engine, the power piston linkage must be strong and robust since it must "support" higher internal pressure (relative to the atmosphere, or buffer space) during parts of the engine cycle. The linkage for the displacer doesn't need to be nearly as strong because the pressure on both sides of it is approximately equal. For an alpha engine, the linkages for both pistons must be strong and robust.
- The crankshaft and flywheel need to be balanced to avoid excessive vibration during high-speed rotation. The easiest way to balance the flywheel is to make sure that its center of rotation coincides with the center of mass. For example, if you can perfectly balance the flywheel about its central axis, then it's balanced. For the crankshaft, the balancing process is a bit trickier. The easiest way to

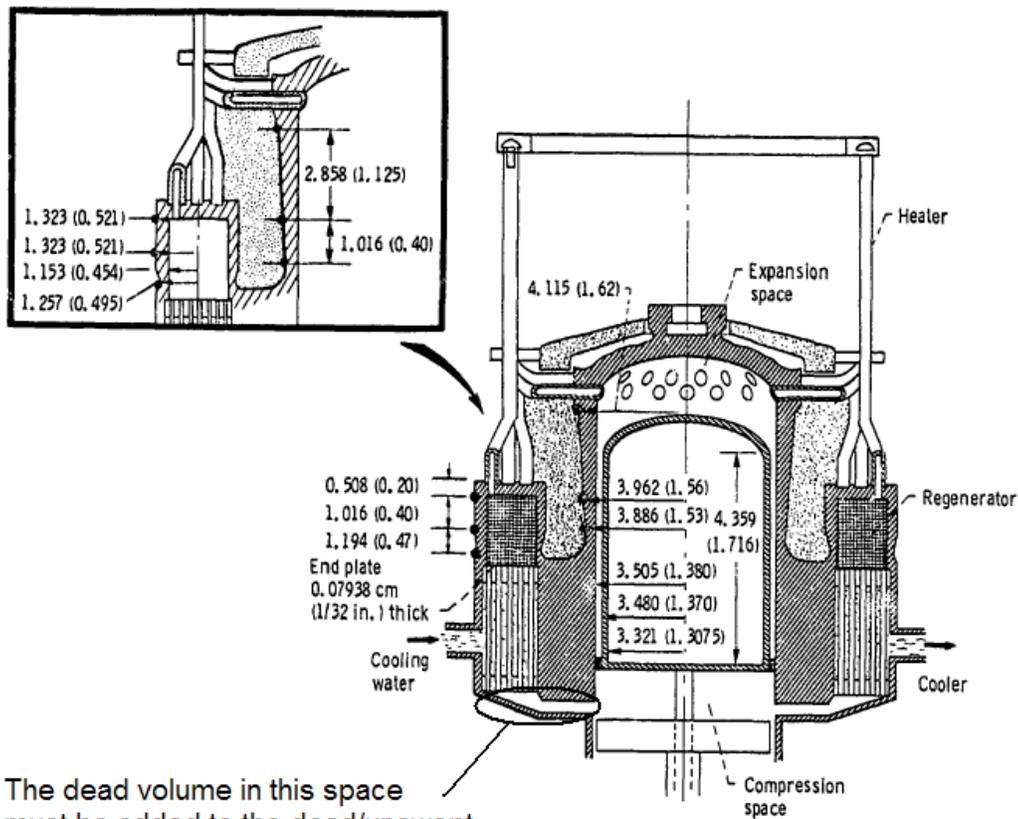
balance a crankshaft is by adding weights to the ends, using trial and error to determine the amount of weight to add, in order to minimize vibration.

- Even though this is not a design manual in the “how to fabricate it” sense, there should be enough information in the 'results' file to figure out how to size the different components for strength. For example, the size of the rods and linkages for the pistons can be determined by looking at the maximum pressure reached inside the engine, and then subtracting the buffer space pressure (if there is a pressurized buffer space). Multiply this pressure difference by the piston area to get the force, and then use a “buckling” equation (from a Machinery’s Handbook) to calculate the size required (maybe using a factor of safety). Similarly, for the displacer (in beta and gamma engines), you must account for the *difference* in pressure between the compression and expansion space (it’s usually small anyway) to determine the force acting on the displacer rod and linkage.

You also have to calculate how thick the walls of your cylinders and heater/cooler tubes need to be to withstand the internal pressure. You also need to consider fatigue stresses since the pressure inside the engine fluctuates. And given that some of the engine components are at high temperature, you will need to look at “creep strength” as well. The details of these considerations are outside the scope of this manual, partly because it’s not the goal of this manual, and also because fabrication guidelines are much more flexible than the information given here, which is based on universal considerations, based on the physics of Stirling engines. However, it may be that you have a “feel” for how big to make the various components based on your own experience, so you won’t need a manual for it. Nonetheless, it’s an important and unavoidable knowledge base to have, especially for those starting out. So it’s likely that some time in the future I will put together a design manual, which describes how to fabricate a Stirling engine. So watch for it.

- When using biomass as heating fuel (such as wood chips or switchgrass), one must be careful of ash buildup on the heater tubes. The airborne ash can melt due to the high temperature, and form an insulating layer on the heater tubes, preventing heat from getting through. This will kill performance. This can be prevented with a gasifier-burner unit, which converts the biomass to combustible gases, which burn cleanly and won’t foul the heat exchanger.
- If water is used as the cooling medium for the cooler, it is typically mixed with antifreeze, especially if the engine is exposed to colder climates.
- If the engine isn’t starting, make sure you are cranking it in the right direction. Remember, in the alpha engine the rotational position of the expansion space piston must be “ahead” of the compression space piston. In the beta and gamma engine the rotational position of the displacer must be “ahead” of the power piston. It may be helpful to draw an arrow indicating the correct direction of rotation.

- A simple way to check the functionality of the engine is to crank the engine by hand or with a motor, in the direction of engine rotation. If the heater tubes get significantly colder and the cooler tubes get significantly warmer this is a sign that there is good energy exchange inside the engine. This interesting phenomenon is a direct consequence of running the engine in “reverse”. By running a Stirling engine in reverse it functions as a refrigerator.
- The picture below is taken from page 16 of the Stirling Engine Design manual. It’s a picture of the GPU-3 engine produced by General Motors in the 1960s (ref: “Stirling Engine Design Manual”, W.R. Martini, Second Edition. The web link is: http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19830022057_1983022057.pdf). Certain parts of the design are worth noting, including the inflow/outflow of the cooling water for the cooler; the long displacer with seals on the compression space side; and the holes at the top of the expansion space, which the heater tubes insert into – this keeps the dead/unswept volume in the expansion space to a minimum. The note I added at the bottom left brings attention to the dead volume between the cooler and compression space. This dead volume is considered part of the dead/unswept volume in the compression space.



The dead volume in this space must be added to the dead/unswept volume in the compression space cylinder

The Radiator and Heat Source (must-read)

As mentioned before, you can use an air-radiator system for the cooler. The coolant flow rate should be high enough so that there is only 10-20 degrees difference between the coolant inlet and outlet temperature. The figure above shows an illustration of cooling water flowing past the cooler tubes. It is necessary to keep the temperature difference small so that the average cooler temperature stays as low as possible. This temperature difference will be the same as the inlet and outlet temperature difference in the radiator. Your radiator must be large enough to accommodate the rate of heat removal from the cooler. The rate of heat removal is based on the temperature difference and the coolant flow rate.

You might be able to use an old car radiator system, but you will have to know its rate of heat removal and the temperature difference it can maintain between the inlet and outlet of the radiator (ideally within a few degrees). Assuming water is the coolant we can estimate the flow rate necessary to maintain a given temperature difference between the inlet and outlet of the cooler. Set T_i as the inlet temperature and T_o as the outlet temperature. The flow rate in liters/second = (Heat removal from cooler, as given in the 'results' file)/{4200*($T_o - T_i$)}. For example, for $T_i = 20$ degrees Celsius, $T_o = 40$ degrees Celsius, and heat removal from cooler = 1340 W, the flow rate = 0.016 liters/second.

The average temperature of the cooler is roughly $(T_o + T_i)/2$. This is the value entered in the input file, under 'Cooler temperature' (in Kelvin).

The size of radiator is such that it can: 1) accommodate the specified flow rate, 2) maintain a temperature difference of $(T_o - T_i)$, and 3) accommodate an inlet temperature of T_i . If you contact a radiator supplier and give them this information they can recommend a suitable size and type of radiator system.

Matching the heating fuel to the specific heat input requirement is more difficult than sizing a radiator for the cooler. One must consider: The type of fuel burned, the fuel flow rate (if applicable), and the efficiency of the air Preheater (if used). If burning biomass or using concentrated solar energy then obviously you cannot change the rate of heat input. So in this case it may be best to just try it out and "see what happens". And if you find that you need more heat energy to maintain the desired heater temperature, then consider switching to another fuel source, perhaps a liquid fuel where you can adjust the flow rate.

Offset Angle in the Expansion and Compression Space Volume (optional)

This wasn't mentioned before to avoid confusion, but an offset angle is used in the volume expressions for the expansion and compression space.

Using the expression given previously for the expansion space (for all three engine configurations), we can introduce an offset angle ϕ

$$V_e = V_{\min,e} + \frac{1}{2} \cdot V_{swept,e} + \frac{1}{2} \cdot V_{swept,e} \cdot \sin(\theta + \phi + \beta)$$

Similarly, for the compression space (for all three engine configurations):

$$V_c = V_{\min,c} + \frac{1}{2} \cdot V_{swept,c} + \frac{1}{2} \cdot V_{swept,c} \cdot \sin(\theta + \phi)$$

This purpose of this angle ϕ is to “shift” V_e and V_c so that, at crank angle $\theta = 0$ (the starting position), the internal engine volume is at maximum value. Maximum engine volume corresponds to the “rest” position when the engine is pressurized against atmospheric pressure. In other words, when pressurizing the engine the pistons move “outwards” until they can’t move any further, and this final position corresponds to maximum engine volume. Consequently this becomes the starting position.

The offset value used is:

$$\phi = \tan^{-1} \left(\frac{V_{swept,c} + V_{swept,e} \cos \beta}{V_{swept,e} \sin \beta} \right)$$

This results in maximum engine volume when $\theta = 0$. The mathematical proof of this is not shown here.

In order for ϕ to be calculated correctly, one must account for the sign of the numerator and denominator, in the above equation. This allows the angle to be calculated based on the location of x, y in the plane, where x = (the denominator) and y = (the numerator). In Fortran, the function used for calculating ϕ is ATAN2(y,x).

The offset angle ϕ will not be mentioned in the mathematical section later on because it is not a necessary part of the mathematical and thermodynamic descriptions. But just keep in mind that it is in fact included in the equations of the program.

Note that the inclusion of ϕ does not affect the solution. It is included simply because it is more realistic from a physical point of view.

Excerpts Taken From Stirling Engine Design Manual (optional)

Below are some main points taken directly from: “Stirling Engine Design Manual”, W. R. Martini, University of Washington, April 1978.

Web link:

http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19780016056_1978016056.pdf

Page 10 – Typically half of the engine volume is in the heat exchangers and ducts, and this reduces the power output to about half of what it is theoretically with no dead volume.

Page 10 – The typical gas cooler is made up of many parallel small diameter tubes with the working gas inside the tubes. The typical regenerator is made up of many layers of stacked mesh screens lightly sintered together. The gas heater is similar to the gas cooler. Combining the heat exchangers with the variable volume spaces is theoretically a good way to reduce dead volume.

Page 10 - Very high gas pressures are used because power density is proportional to average gas pressure. Losses increase only slowly with gas pressure. Higher heater temperature also increases power density and adds to the thermal efficiency.

Page 11 – The *Rinia arrangement* is an engine design sometimes used. This is a design using multiple cylinders.

Page 11 – Piston seals cannot be lubricated with oil, as this will quickly foul the heat exchangers. Filled Teflon piston rings are usually used. Specially designed mechanical seals or oil backed roll sock seals are used to almost eliminate leakage. In some low power systems, diaphragms or bellows are used instead of pistons, which eliminates sliding friction, and mechanical wear.

Page 14 – Hydrogen and helium as the working gas, works much better than any other gas (especially at high pressures). Hydrogen is best because it has the highest thermal conductivity, the lowest viscosity and a low heat capacity on a volume basis. Only a small amount of heat is needed to change its temperature. However, hydrogen permeates through metals and no container is completely impermeable. And hydrogen is flammable. Some metals are embrittled by hydrogen. Helium, however, is inert and can be permanently contained in metal. It has an even lower volumetric heat capacity than hydrogen but its viscosity is twice that of hydrogen. It has been shown that hydrogen, helium, or nitrogen can all power a Stirling engine with the same thermal efficiency. However, the heater, cooler, and regenerator would be designed quite differently for each. Hydrogen allows for a broader range of high thermal efficiency operation than helium.

Page 52 – There are three basic analysis methods for Stirling engines: First order (Schmidt), Second Order (Schmidt or similar, with fluid friction and mechanical losses factored in), Third Order (discretizing the working space into nodes, using fluid momentum equations, heat transfer equations, and energy equations) – very few assumptions, and more accurate. One assumption not necessarily made in the Third Order analysis is that pressure is the same everywhere inside the working spaces. Pressure

gradients do exist and time steps have to be chosen to be smaller than the time required for the speed of sound to travel across a gradient.

Page 68 – Using the Schmidt analysis (isothermal space assumption), there is a 1-2% difference in work prediction from using a more realistic adiabatic space analysis. This comparison is based on the usual dead volume to total volume ratios used in Stirling engines.

Page 110 – Practical engines often have short cranks that lead to piston motions that vary quite differently from sinusoidal motion (as assumed by the Schmidt analysis).

Page 117 – Using a Second Order approximation, fluid friction (flow) losses should be less than 10% of basic power (within the desired range of operating speed). One can use empirical correlations for fluid flow through porous media and tubes, to estimate flow friction. These correlations are for steady, fully developed flow, so it is only an approximation for fluid that is moving back and forth through heat exchangers and regenerator. Also, the mass flow into and out of the regenerator is not quite in phase due to accumulation and depletion of mass in the regenerator. And note that the mass flow at the cold end is much more than the mass flow at the hot end due to the higher gas density. The average mass flow rate and the average fraction of the total cycle time that gas is flowing in one direction at the hot end of the regenerator is used for the heater flow friction and heat transfer calculations. The average mass flow rate and the average fraction of the total cycle time that gas is flowing in one direction at the cold end of the regenerator is used for the cooler flow friction and heat transfer calculations. For the regenerator, the mean of the above two flows and of the above two fractions will be used.

Page 121 – The so-called *windage loss* is the total pumping (flow) loss per cycle as the gas passes through the heater, cooler and regenerator. During each cycle, the working gas flows twice through these parts.

Page 122 – the so-called *reheat loss* is caused by regenerator inefficiency (non-ideal).

Page 143 to 152 – For Third Order methods there are four fundamental equations that must be solved: continuity, momentum, energy, and equation of state. More specifically these are (respectively): conservation of mass, Navier-Stokes fluid momentum equations, energy balance, and $p(V) = m(T)$ using normalizing parameters (Urieli).

Sensitivity Analysis (must-read)

A sensitivity analysis was done to determine how pressure drop across the regenerator affects engine performance. In the model used, it is assumed that the percentage of pressure drop across the regenerator is 80% of the total pressure drop across the heater, regenerator, and cooler. In reality this percentage can vary between 70-90% according to reference [3]. However, the range of 70-90% (although somewhat small) actually represents a large range of pumping (flow) losses that can exist inside the engine. For example, if the pumping loss in the heater and cooler is 100 W, and the pumping loss in

the regenerator is 70% of the total, this means that the total pumping loss is $100/(1-0.70) = 333$ W. And if the pumping loss in the regenerator is 90% of the total, this means that the total pumping loss is $100/(1-0.90) = 1000$ W – much higher! And with 80% pressure drop the total pumping loss is somewhere in between. Therefore, it is very informative to do a sensitivity analysis to see how the results are affected. Furthermore, covering a large range of pumping loss also helps compensate for the sensitivity of the solution to the specific pumping loss in the heater and cooler. For instance, with 100 W of pumping loss in the heater and cooler there is a range of total pumping loss of 333-1000 W. But if (for example) the pumping loss in the heater and cooler were 120 W, the range of total pumping loss would be 400-1200 W. This range is quite a bit larger towards the high end. But as it turns out, it will not make much difference in terms of the optimal number of heater and cooler tubes, and regenerator void volume (for maximum power), as you shall see.

In a nutshell, the sensitivity analysis I did shows that the optimal number of tubes (with proportional regenerator void volume) resulting in maximum power is barely affected by the range of 70-90%. This is a very useful result. It says that engine optimization is almost entirely independent of the assumed percentage of pressure drop in the regenerator! So the optimal number of tubes, assuming 80% pressure drop, will be the same as if you had 70% or 90% pressure drop (and any pressure drop in between). And by extension, a pressure drop outside this range, by a few percent, would probably still not significantly change the optimal number of tubes.

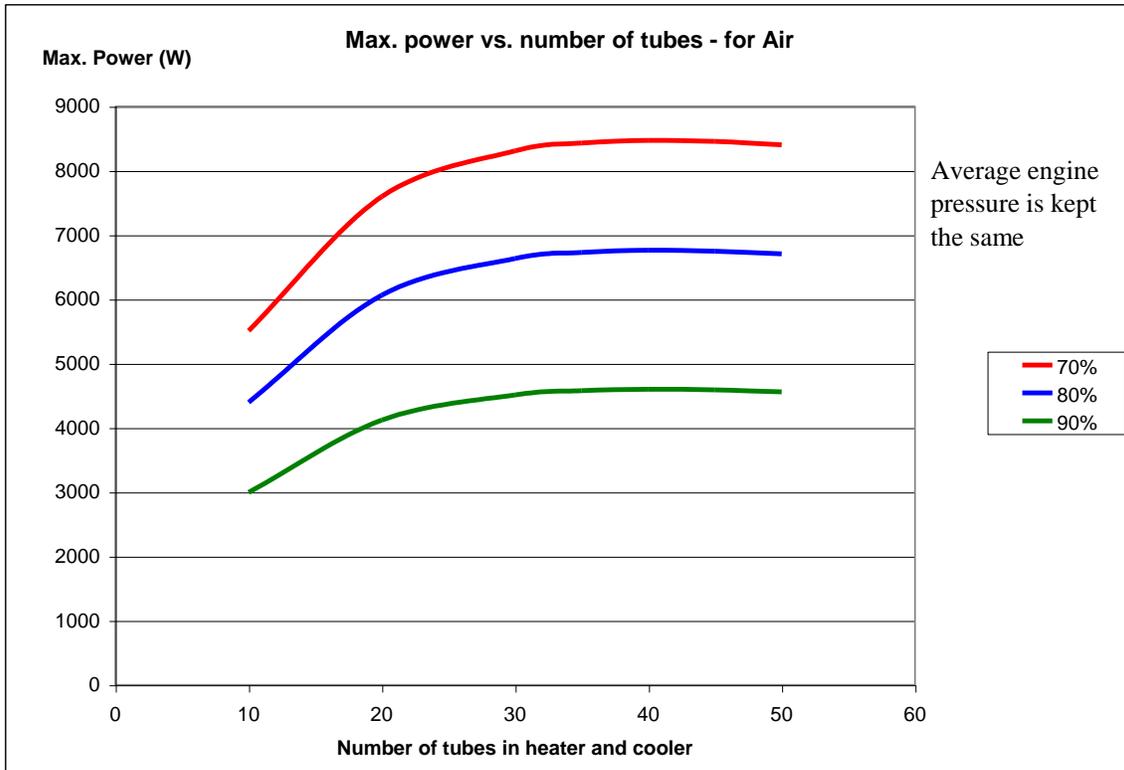
Another interesting result of the sensitivity analysis is that the thermal efficiency at maximum power is barely affected by the assumed percentage of pressure drop across the regenerator.

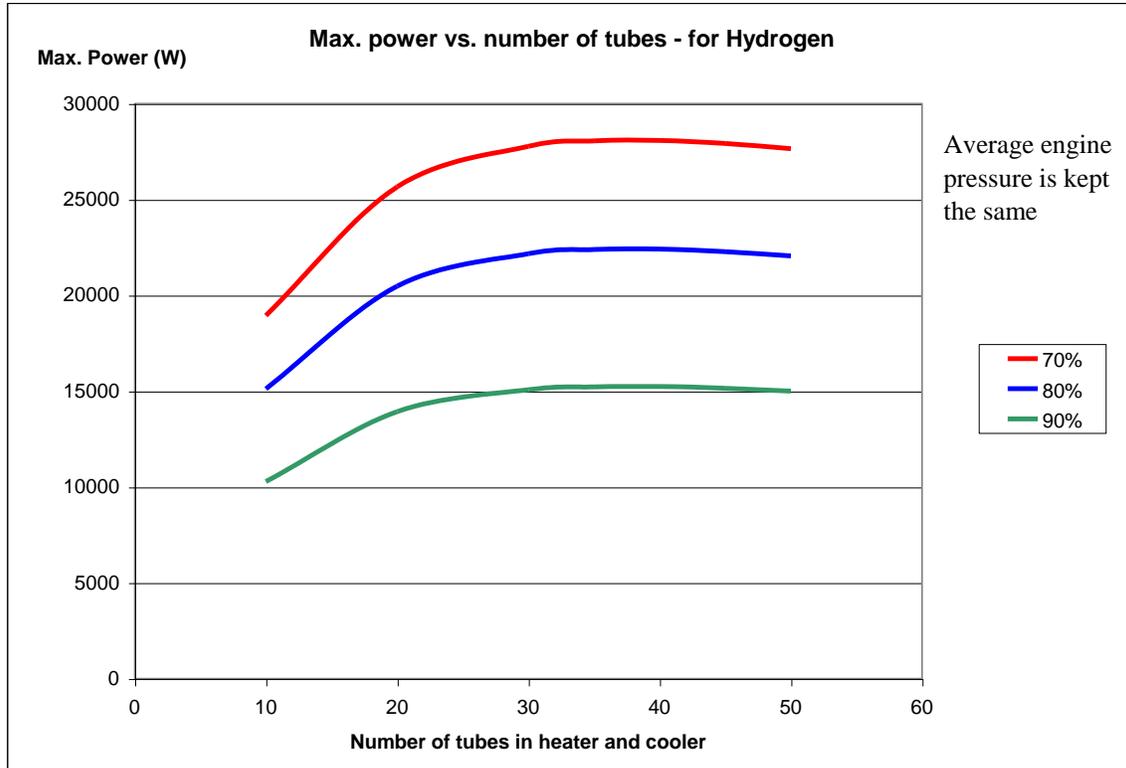
But on the other hand, things like engine speed and the specific value of maximum power are greatly affected by the range 70-90%. But correctly predicting power and speed is perhaps not as important as correctly predicting the optimal number of tubes and regenerator void volume, for maximum power. But it is still good to have a sense of the range of power and operating speed anyway, given the possible range of regenerator pressure drop. This will be shown in the following graphs.

The sensitivity analysis was done for an example engine configuration, using the input data shown in Appendix G. Hydrogen and air was used as the working gas. I chose to use these two gases because they approximately represent the two extremes of the working gases you can use, in terms of properties. Most gases fall somewhere in between these two, in terms of their properties. So it is easier to generalize the results based on the results for air and hydrogen.

The following two graphs show the maximum power, versus number of tubes, with air and hydrogen as the working gas. The results are given with 70, 80, and 90% pressure drop across the regenerator. One can clearly see that the optimal number of tubes is about 40. For all three curves there is very little difference in maximum power for the (tube

number) range 35 to 45. In other words, there is almost a perfectly flat plateau in this range, for all three curves.





Note that by adjusting the scale factor in the input file the average engine pressure was kept the same for all the curves. This was done to ensure consistency. Since engine power is directly proportional to average engine pressure, one can only compare engine power between two configurations when the average engine pressure is the same.

Notice that the maximum power for hydrogen is much higher than the maximum power for air. This is because hydrogen has a lower viscosity and much lower density than air, resulting in much lower pumping (flow) losses.

As a rough estimate, you can approximate the percentage variation in maximum engine power relative to the 80% pressure drop case. One can approximate the percentage variation (for the general case) from the curves shown above. Looking at the curves this variation is about +25%/-30%. So if the program predicts a power of 6800 W, the actual power may lie between 4800 and 8500 W. In addition, the heat input requirement (for the heater), and heat removal requirement (for the cooler) also varies by approximately +25%/-30%.

Given that this percentage variation is the same for air and hydrogen, you can say that it is roughly the same for other working gases as well.