Pulse Tube Cryocoolers for Cooling Infrared Sensors*

Ray Radebaugh

National Institute of Standards and Technology
Boulder, Colorado 80303 USA

ABSTRACT

This paper reviews recent advances in pulse tube cryocoolers and their application for cooling infrared sensors. There are many advantages of pulse tube cryocoolers over Stirling cryocoolers associated with the absence of moving parts in the cold head. Efficiencies have been improved considerably in the last few years to where they equal or even exceed the efficiencies of Stirling cryocoolers. The use of inertance effects and double inlets to improve the efficiencies will be discussed. Pulse tube cryocoolers are now being used or considered for use in cooling infrared detectors for many space applications. One disadvantage of pulse tube coolers is the difficulty in scaling them down to sizes as small as 0.15 W at 80 K while maintaining high efficiency. A second disadvantage is the larger diameter cold finger required for the same refrigeration power because of the presence of the pulse tube. These two disadvantages have limited their use so far in cooling infrared sensors for many military tactical applications. Progress in overcoming these disadvantages is discussed.

Keywords: Cryocoolers, Cryogenics, Detectors, Infrared, Night vision, Pulse tubes, Refrigerators, Review, Sensors, Stirling

1. INTRODUCTION

Small cryogenic refrigerators (cryocoolers) are required for a wide variety of applications, and the number of applications keeps expanding as improvements to cryocoolers are made. One of the earliest applications, and one that appeared about 50 years ago, was for cooling infrared sensors to about 80 K for night vision capability of the military. Over 125,000 Stirling cryocoolers for this tactical military application have been produced to date. Refrigeration powers vary from about 0.15 W to 1.75 W. In the last decade or so the desire for night vision surveillance and missile detection from satellites has prompted much research on improved cryocoolers to meet the very stringent requirements for space use. Each application has a particular set of requirements that have led to many recent improvements in cryocoolers. Still, some of the problems associated with cryocoolers have hampered the marketing of many potential applications. The major problems associated with cryocoolers are listed in Table 1. For satellite applications all but the last problem in Table 1 are of major concern, even though a space-qualified cryocooler can cost $1.0M or more. Lifetimes of 10 years are now standard requirements for most space applications. Until recently efficiencies of 3 to 10% of Carnot were typical of these small cryocoolers, whereas now efficiencies of 15 to 25% of Carnot are now possible.

Table 1. Cryocooler problems.

- Reliability
- Efficiency
- Size and weight
- Vibration
- Electromagnetic Interference (EMI)
- Heat rejection
- Cost

* Contribution of NIST, not subject to copyright in the U. S.
New commercial applications, particularly the use of high temperature superconductors as microwave filters in base stations for wireless communication systems, have stimulated much research on ways to reduce the cost of cryocoolers while still maintaining a lifetime goal of 3 to 5 years. Research on various types of cryocoolers has been carried out with the goal of meeting the requirements for a particular application. As a result, improvements have been made in all the types of cryocoolers. This paper briefly reviews some of the recent advances in pulse tube cryocoolers. More extensive reviews of all cryocoolers as well as pulse tube cryocoolers have been given by the author. The advantages of pulse tube cryocoolers make them natural candidates for the cooling of infrared detectors for both tactical and space applications. Pulse tube cryocoolers are now being developed for many space applications, but the tactical applications have had a particular set of historical requirements that have made it difficult to use pulse tube cryocoolers in place of the Stirling cryocoolers. This paper also discusses some of these problems and how recent advances may make it possible to meet these requirements, at least for larger systems.

2. TYPES OF CRYOCOOLERS

Though the focus of this paper is on pulse tube cryocoolers, a brief introduction of all the cryocooler types is given here because some comparisons will be given later that are more meaningful if there is some understanding of the various types. Figure 1 shows the five types of cryocoolers in common use today. The Joule-Thomson (JT) and the Brayton cryocoolers are of the recuperative type in which the working fluid flows steadily in one direction, with steady low- and high-pressure lines, analogous to DC electrical systems. The compressor has inlet and outlet valves to maintain the steady flow. The recuperative heat exchangers transfer heat from one flow stream to the other over some distance. Recuperative heat exchangers with the high effectiveness needed for cryocoolers can be expensive to fabricate. The three regenerative cryocoolers shown in Fig. 1 operate with an oscillating flow and an oscillating pressure, analogous to AC electrical systems. Frequencies vary from about 1 Hz for the Gifford-McMahon (GM) and some pulse tube cryocoolers to about 60 Hz for Stirling and some pulse tube refrigerators.

2.1 Recuperative Cryocoolers

The steady pressure and the steady flow of gas in these cryocoolers allow them to use large gas volumes anywhere in the system with little adverse effects except for larger radiation heat leaks if the additional volume is at the cold end. Thus, it is possible to “pipe cold” to any number of distant locations after the gas has expanded and cooled. In addition, the cold end can be separated from the compressor by a large distance and greatly reduce the electromagnetic interference (EMI) and vibration associated with the compressor. Oil removal equipment with its large gas volume can also be incorporated in these cryocoolers at the warm end of the heat exchanger to remove any traces of oil from the working gas before it is cooled in the heat exchanger. Unlike conventional refrigerators operating near ambient temperature, any oil in the working fluid will freeze at cryogenic temperatures and plug the system.

![Schematics of five common types of cryocoolers.](image-url)
2.1.1 Joule-Thomson cryocoolers

The Joule-Thomson cryocoolers produce cooling when the high-pressure gas expands through a flow impedance (orifice, valve, capillary, porous plug), often referred to as a JT valve. The expansion occurs with no heat input or production of work, thus, the process occurs at a constant enthalpy. The heat input occurs after the expansion and is used to warm up the cold gas or to evaporate any liquid formed in the expansion process. In an ideal gas the enthalpy is independent of pressure for a constant temperature, but real gases experience an enthalpy change with pressure. Thus, cooling in a JT expansion occurs only with real gases and at temperatures below the inversion curve. In fact, for a given pressure change the amount of cooling increases as the temperature is lowered and reaches a maximum around the critical point. Typically, nitrogen or argon is used in JT coolers, but pressures of 20 MPa (200 bar) or more on the high-pressure side are needed to achieve reasonable cooling. Such high pressures are difficult to achieve and require special compressors with limited lifetimes.

The main advantage of JT cryocoolers is the fact that there are no moving parts at the cold end. The cold end can be miniaturized and provide a very rapid cool-down. This rapid cool-down (a few seconds to reach 77 K) has made them the cooler of choice for cooling infrared sensors used in missile guidance systems. These coolers utilize a small cylinder pressurized to about 45 MPa with nitrogen or argon as the source of high-pressure gas. In this open-cycle mode, cooling lasts for only a few minutes until the gas is depleted. Figure 2 shows a typical JT cryocooler used for missile guidance. Miniature finned tubing is used for the heat exchanger. An explosive valve is used to start the flow of gas from the high-pressure bottle, and after flowing through the system, the gas is vented to the atmosphere.

A disadvantage of the JT cryocooler is the susceptibility to plugging by moisture of the very small orifice (possibly experienced in a test of the long-range U. S. Missile Defense System in January, 2000). Another disadvantage is the low efficiency when used in a closed cycle mode. Compressor efficiencies are very low when compressing to such high pressures.

Recent advances in JT cryocoolers have been associated with the use of mixed-gases as the working fluid rather than pure gases. The use of mixed gases was first proposed in 1936 for the liquefaction of natural gas, but it was not used extensively for this purpose until the last 20 or 30 years. It is commonly referred to as the mixed refrigerant cascade (MRC) cycle. The use of small JT coolers with mixed gases for cooling infrared sensors was first developed under classified programs in the Soviet Union during the 1970s and 1980s. Such work was first discussed in the open literature by Little4. Missimer5 and Radebaugh2,6 review the use of mixed gases in JT cryocoolers. Typically, higher boiling-point components, such as methane, ethane, and propane can be added to nitrogen to make the mixture behave more like a real gas over the entire temperature range. Enthalpy changes as much as 5 times that with pure nitrogen are possible with pressures of only 2.5 MPa. Such large enthalpy changes then lead to greatly increased efficiencies with the JT cryocooler at pressures that can be achieved in conventional compressors used for domestic or commercial refrigeration. The higher boiling point components must remain a liquid at the lowest temperature. In general, the freezing point of a mixture is less than that of the pure fluids, so temperatures of 77 K are possible with the nitrogen-hydrocarbon mixtures even though the hydrocarbons freeze in the range of 85 to 91 K as pure components. The presence of propane also increases the solubility of oil in the mixture at 77 K so that less care is needed in removing oil from the mixture when using an oil-lubricated compressor. Much research is currently underway pertaining to the solubility of oil in various mixtures and the freezing point of mixtures. Marquardt et al.7 discuss the optimization of gas mixtures for a given temperature range and show how a mixed-gas JT cryocooler for missile guidance.
cryocooler can be used for a cryogenic catheter only 3 mm in diameter. Such miniature systems could also be used for cooling infrared sensors.

2.1.2 Brayton cryocoolers

In Brayton cryocoolers (sometimes referred to as the reverse-Brayton cycle to distinguish it from a heat engine) cooling occurs as the expanding gas does work. Figure 1 shows a reciprocating expansion engine for this purpose, but an expansion turbine supported on gas bearings is more commonly used to give high reliability. According to the First Law of Thermodynamics the heat absorbed with an ideal gas in the Brayton cycle is equal to the work produced. This process is then more efficient than the JT cycle and it does not require as high a pressure ratio. The Brayton cycle is commonly used in large liquefaction plants. For small Brayton cryocoolers the challenge is fabricating miniature turboexpanders that maintain high expansion efficiency. Turbine diameters of about 6 mm on shafts of 3 mm diameter are typical in systems reviewed by McCormick et al. for use in space applications of cooled infrared sensors. Turbine speeds of 2000 to 5000 rev/s are typical. Centrifugal compressors providing a pressure ratio of about 1.6 with a low-side pressure of 0.1 MPa are used with these systems. A similar system to be used on the Hubble Space Telescope will provide 7 W of cooling at 70 K. The working fluid used in the turbo-Brayton cryocoolers is usually neon when operating above 35 K, but helium is required for lower temperatures.

An advantage of the Brayton cryocooler is the very low vibration associated with rotating parts in a system with turboexpanders and centrifugal compressors. This low vibration is often required with sensitive telescopes in satellite applications, such as with the Hubble telescope. The expansion engine provides for good efficiency over a wide temperature range, although not as high as some Stirling and pulse tube cryocoolers at temperatures above about 50 K. The low-pressure operation of the miniature Brayton systems requires relatively large and expensive heat exchangers.

2.2 Regenerative Cryocoolers

These cryocoolers operate with oscillating pressures and mass flows in the cold head. The oscillating pressure can be generated with a valveless compressor (pressure oscillator) as shown in Fig. 1 for the Stirling and pulse tube cryocoolers, or with valves that switch the cold head between a low and high pressure source, as shown for the Gifford-McMahon cryocooler. In the latter case a conventional compressor with inlet and outlet valves is used to generate the high- and low-pressure sources. An oil-lubricated compressor is usually used and oil removal equipment can be placed in the high-pressure line where there is no pressure oscillation. The use of valves greatly reduces the efficiency of the system. Pulse tube cryocoolers can use either source of pressure oscillations, even though Fig. 1 indicates the use of a valveless compressor. The valved compressors are modified air conditioning compressors, and they are used primarily for commercial applications where low cost is very important. The main heat exchanger in regenerative cycles is called a regenerator. In a regenerator, incoming hot gas transfers heat to the matrix of the regenerator, where the heat is stored for a half cycle in the heat capacity of the matrix. In the second half of the cycle the returning cold gas, flowing in the opposite direction through the same channel, absorbs heat from the matrix and returns the matrix to its original temperature before the cycle is repeated. Very high surface areas for enhanced heat transfer are easily achieved in regenerators through the use of stacked fine-mesh screen or packed spheres. Gifford-McMahon cryocoolers are seldom used in cooling infrared sensors because of their low efficiency and large size. Their input powers are usually 1 to 6 kW. Therefore, they will not be discussed further.

2.2.1 Stirling cryocoolers

The Stirling cycle, as invented and patented by Robert Stirling in 1815, was first used as a prime mover. In 1834 John Herschel proposed its use as a refrigerator in producing ice. It was not until about 1861 that Alexander Kirk reduced the concept to practice. Air was used as the working fluid in these early regenerative systems. Very little development of Stirling refrigerators occurred until 1946 when a Stirling engine at a Dutch company was run in reverse with a motor and was found to liquefy air on the cold tip. The engine used helium as the working fluid, since earlier work at the company showed helium to give much improved performance to the engines. Stirling cryocoolers have been used for about the last
40 years in cooling infrared sensors for tactical military applications in such equipment as tanks and airplanes. They cannot provide the very fast cooldown times of JT cryocoolers, so they are not used on missiles for guidance. The long history of the Stirling cryocooler in cooling infrared equipment has resulted in many specifications being tailored to the geometry characteristics of the Stirling cryocooler. As a result, newer cryocoolers, like the pulse tube cryocooler, with different geometries are difficult to adapt to the geometry specifications.

A pressure oscillation by itself in a system would simply cause the temperature to oscillate and produce no refrigeration. In the Stirling cryocooler the second moving component, the displacer, is required to separate the heating and cooling effects by causing motion of the gas in the proper phase relationship with the pressure oscillation. When the displacer in Figure 1c is moved downward, the helium gas is displaced to the warm end of the system through the regenerator. The piston in the compressor then compresses the gas, and the heat of compression is removed by heat exchange with the ambient. Next the displacer is moved up to displace the gas through the regenerator to the cold end of the system. The piston then expands the gas, now located at the cold end, and the cooled gas absorbs heat from the system it is cooling before the displacer forces the gas back to the warm end through the regenerator. There is little pressure difference across the displacer (only enough to overcome the pressure drop in the regenerator) but there is a large temperature difference. Most actual Stirling cryocoolers have the regenerator inside the displacer instead of external as shown in Fig. 1c. The resulting single cylinder provides a convenient cold finger.

In practice, motion of the piston and the displacer are nearly sinusoidal. The correct phasing occurs when the volume variation in the cold expansion space leads the volume variation in the warm compression space by about 90°. With this condition the mass flow or volume flow through the regenerator is approximately in phase with the pressure. In analogy with AC electrical systems, real power flows only when current and voltage are in phase with each other. Without the displacer in the Stirling cycle the mass flow leads the pressure by 90° and no refrigeration occurs. Though the moving piston causes both compression and expansion of the gas, net power input is required to drive the system. The moving displacer reversibly extracts net work from the gas at the cold end and transmits it to the warm end where it contributes some to the compression work. In an ideal system, with isothermal compression and expansion and a perfect regenerator, the process is reversible. Thus, the coefficient of performance $COP$ for the ideal Stirling refrigerator is the same as the Carnot $COP$ given by

$$COP_{Carnot} = \frac{\dot{Q}_c}{W_0} = \frac{T_c}{T_h - T_c},$$

where $\dot{Q}_c$ is the net refrigeration power, $W_0$ is the power input, $T_c$ is the cold temperature, and $T_h$ is the hot temperature. The occurrence of $T_c$ in the denominator arises from the PV power (proportional to $T_c$) recovered by the expansion process and used to help with the compression. Practical cryocoolers have $COP$ values that range from about 1 to 25% of the Carnot value.

Figure 3 shows the four sizes of Stirling cryocoolers that are currently used for military tactical applications. The refrigeration powers listed for each cooler are for a temperature of about 77 to 80 K, except the 1.75 W system, which is for a temperature of 67 K. Their specified minimum efficiencies range from about 3 to 6% of Carnot as the size increases. All of the coolers shown in Fig. 3 use linear drive motors with a dual-opposed arrangement to reduce vibration. The linear drive reduces side forces between the piston and the cylinder and the Mean-Time-To-Failure (MTTF) is at least 4000 hours. The displacer is driven pneumatically with the oscillating pressure in the system and because there is only one displacer it gives rise to considerable vibration. Efforts are currently underway to increase the MTTF of these Stirling cryocoolers since they are the least reliable component in an infrared system. The development of cryocoolers for space applications has led to greatly improved reliabilities, and a MTTF of 10 years is now usually specified for these applications. The Stirling cooler was first used in these space applications after flexure bearings were developed for supporting the piston and displacers in their respective cylinders with little or no contact in a clearance gap of about 15 µm. Figure 4 shows two examples of flexure bearing geometries used in a compressor. The displacer would use the same type of flexure. These flexure-supported Stirling cryocoolers were initially very expensive, but advances in manufacturing have brought down the price to where these flexures are now being investigated for use in compressors in tactical
applications. Such compressors would also be useful for pulse tube cryocoolers, since they can use the same type of compressor.

2.2.2 Pulse tube cryocoolers

The moving displacer in the Stirling and Gifford-McMahon cryocoolers has several disadvantages. It is a source of vibration, has a limited lifetime, and contributes to axial heat conduction as well as to a shuttle heat loss. In the pulse tube cryocooler, shown in Figure 1d, the displacer is eliminated. The proper gas motion in phase with the pressure is achieved by the use of an orifice, along with a reservoir volume to store the gas during a half cycle. The reservoir volume is large enough that negligible pressure oscillation occurs in it during the oscillating flow. The oscillating flow through the orifice separates the heating and cooling effects just as the displacer does for the Stirling and Gifford-McMahon refrigerators. The orifice pulse tube refrigerator (OPTR) operates ideally with adiabatic compression and expansion in the pulse tube. Thus, for a given frequency there is a lower limit on the diameter of the pulse tube in order to maintain

Figure 3. Four sizes of Stirling cryocoolers with dual-opposed linear compressors.

Figure 4. Schematic of (a) linear compressor with (b) two types of flexure bearings.
adiabatic processes. The four steps in the cycle are as follows. (1) The piston moves down to compress the gas (helium) in the pulse tube. (2) Because this heated, compressed gas is at a higher pressure than the average in the reservoir, it flows through the orifice into the reservoir and exchanges heat with the ambient through the heat exchanger at the warm end of the pulse tube. The flow stops when the pressure in the pulse tube is reduced to the average pressure. (3) The piston moves up and expands the gas adiabatically in the pulse tube. (4) This cold, low-pressure gas in the pulse tube is forced toward the cold end by the gas flow from the reservoir into the pulse tube through the orifice. As the cold gas flows through the heat exchanger at the cold end of the pulse tube it picks up heat from the object being cooled. The flow stops when the pressure in the pulse tube increases to the average pressure. The cycle then repeats. The function of the regenerator is the same as in the Stirling and Gifford-McMahon refrigerators in that it precools the incoming high-pressure gas before it reaches the cold end.

One function of the pulse tube is to insulate the processes at its two ends. That is, it must be large enough that gas flowing from the warm end traverses only part way through the pulse tube before flow is reversed. Likewise, flow in from the cold end never reaches the warm end. Gas in the middle portion of the pulse tube never leaves the pulse tube and forms a temperature gradient that insulates the two ends. Roughly speaking, the gas in the pulse tube is divided into three segments, with the middle segment acting like a displacer but consisting of gas rather than a solid material. For this gas plug to effectively insulate the two ends of the pulse tube, turbulence in the pulse tube must be minimized. Thus, flow straightening at the two ends is crucial to the successful operation of the pulse tube refrigerator. The pulse tube is the unique component in this refrigerator that appears not to have been used previously in any other system. It could not be any simpler from a mechanical standpoint. It is simply an open tube. But the thermohydrodynamics of the processes involved in it are extremely complex and still not well understood or modeled. The overall function of the pulse tube is to transmit hydrodynamic or acoustic power in an oscillating gas system from one end to the other across a temperature gradient with a minimum of power dissipation and entropy generation.

Pulse tube refrigerators were invented by Gifford and Longsworth in the mid 1960s, but that type was different than what is shown in Fig. 1d and only reached a low temperature of 124 K. In 1984 Mikulin et al. introduced the concept of an orifice to the original pulse tube concept and reached 105 K. In 1985 Radebaugh et al. changed the location of the orifice to that shown in Fig. 1d and reached 60 K. Further improvements since then have led to a low temperature limit of about 20 K with one stage and 2 K with two stages. (See discussion in Reference 3.) There are three different geometries that have been used with pulse tube cryocoolers as shown in Fig. 5. The inline arrangement is the most efficient because it requires no void space at the cold end to reverse the flow direction nor does it introduce turbulence into the pulse tube from the flow reversal. The disadvantage is the possible awkwardness associated with having the cold

Figure 5. Three different geometries for pulse tube cryocoolers.
plate located between the two warm ends. The most compact arrangement and the one most like the geometry of the Stirling cryocooler is the coaxial arrangement. That geometry has the potential problem of a mismatch of temperature profiles in the regenerator and in the pulse tube that would lead to steady heat flow between the two components and a reduced efficiency. However, that problem has been minimized, and a coaxial geometry has been used as an oxygen liquefier for NASA with an efficiency of 17% of Carnot\textsuperscript{18}. The coaxial geometry is the only geometry that could possibly meet the geometry specifications of the Standard Advanced Dewar Assembly (SADA) for second-generation thermal imaging systems\textsuperscript{14, 19} of the military. The absence of a moving displacer in pulse tube cryocoolers gives them many potential advantages over Stirling cryocoolers for the cooling of infrared sensors. Early pulse tube cryocoolers were not nearly as efficient as Stirling cryocoolers, but advances in the last ten years have brought pulse tube refrigerators to the point of being the most efficient of all cryocoolers. Some details of this rapid progress are given in the following section.

3. ADVANCES IN PULSE TUBE CRYOCOOLERS

3.1. Phase Shift Mechanisms

In regenerative cryocoolers the optimum phase relationship between the pressure and the mass (or volume) flow is to have the two in phase with each other near the center of the regenerator. Because of the gas volume in the regenerator the mass flow at the warm end of the regenerator will lead the pressure, but the flow at the cold end will lag the pressure. With such a phase relationship the magnitude of the average mass flow rate through the regenerator for a given PV power is minimized. Because the regenerator losses (such as pressure drop and imperfect heat transfer) depend mostly on the magnitude of the mass flow rate, the regenerator losses are minimized and the system efficiency is maximized. In a Stirling cryocooler, where the displacer is driven, the optimum phase angle can always be achieved. In the orifice pulse tube refrigerator the mass flow rate and the pressure are in phase at the orifice because of the purely resistive nature of the flow impedance in an orifice. In practice the orifice can be any resistive flow element such as an orifice, a needle valve (for adjustment purposes), a capillary, or a porous plug. In all cases the flow through the element is in phase with the dynamic pressure (same as the pressure drop because the pressure in the reservoir is constant). Because of the gas volume in the pulse tube the conservation of mass dictates that the mass flow at the cold end of the pulse tube and regenerator will lead the mass flow at the warm end of the pulse tube (typically about 30° with an optimum pulse tube volume). Consequently, the mass flow at the cold end also leads the pressure instead of lagging the pressure at that location for the minimum regenerator loss. The pressure does not change phase or amplitude through a typical pulse tube, unless frequencies much above 60 Hz are used. The mass flow at the warm end of the regenerator then greatly leads the pressure (perhaps by as much as 60°), which gives rise to a large amplitude of mass flow and causes large regenerator losses for a given PV power flow. Thus, early orifice pulse tube refrigerators were not as efficient as Stirling cryocoolers, and they required a compressor with a larger swept volume. Then mechanisms were found to shift the phase without the need for a moving part.

3.1.1. Double inlet (secondary orifice)

In 1990 Zhu, Wu, and Chen\textsuperscript{20} introduced the concept of a secondary orifice to the OPTR in which the secondary orifice allows a small fraction (about 10%) of the gas to travel directly between the compressor and the warm end of the pulse tube, thereby bypassing the regenerator. They called this the double-inlet pulse tube refrigerator. This bypass flow is used to compress and expand the portion of the gas in the warm end of the pulse tube that always remains at the warm temperature. The bypass flow reduces the flow through the regenerator, thereby reducing the regenerator loss. The flow through the secondary orifice is in phase with the pressure drop across the regenerator, which, in turn, is approximately in phase with the flow through the regenerator, averaged over its length. Since the flow usually leads the pressure, so too will the flow through the secondary. This secondary flow then forces the flow at the warm end of the pulse tube to lag the pressure, since the sum of the flows must be in phase with the pressure at the primary orifice. The location of the secondary orifice is shown in Fig. 6, along with that of an inertance tube described in the next section.
The double-inlet pulse tube refrigerator led to increased efficiencies, particularly at high frequencies where the regenerator losses would be quite high in a simple OPTR. The secondary orifice was incorporated in the mini pulse tube described by Chan et al.\textsuperscript{21}. This refrigerator produced 0.5 W of cooling at 80 K with only 17 W of input power. About 30 of these mini pulse tube refrigerators have been built and are scheduled for a variety of space missions, mostly for cooling infrared detectors. The secondary orifice has been used on a majority of small pulse tube refrigerators built since about 1991.

3.1.2. DC flow or streaming

Though the introduction of the secondary orifice usually led to increased efficiencies compared to the OPTR, it also introduced a problem. Performance of the double inlet pulse tube refrigerator was not always reproducible, and sometimes the cold end temperature would slowly oscillate by several degrees with periods of several minutes or more. Researchers began attributing this erratic behavior to DC flow that can occur around the loop formed by the regenerator, pulse tube, and secondary orifice. Asymmetric flow impedance in the secondary can cause such a DC flow. Even small DC flow carries a large enthalpy flow from the warm to the cold end because of the large temperature difference at the two ends. In 1997 Gedeon\textsuperscript{22} showed that the acoustic power flowing from the warm to the cold end of the regenerator brings about an intrinsic driving force for DC flow or streaming in the same direction as the acoustic power flow. As a result, an asymmetric secondary is required to cancel the intrinsic tendency for this DC flow. In the author’s lab the use of a needle valve for the secondary with the needle pointing toward the warm end of the pulse tube resulted in a no-load temperature on a small pulse tube refrigerator of about 35 K. When the needle valve was reversed, the no-load temperature increased to about 50 K. A tapered tube, also known as a jet pump, has also been used to cancel the DC flow\textsuperscript{23}.

Direct measurements of this small DC flow superimposed on the AC flow would be nearly impossible and have never been attempted. Instead, it is customary to measure the temperature profile on the outside of the regenerator and compare that with the calculated profile or the measured profile when the secondary is closed. Except for very low temperatures the normalized temperature at the midpoint on the regenerator is about 50 to 55% of the total temperature difference. DC flow from the warm end of the regenerator increases the midpoint temperature, whereas flow in the opposite direction reduces the midpoint temperature.

3.1.3. Inertance tube

The conservation of momentum equation for the working fluid is

\[
\frac{\partial \rho \mathbf{v}}{\partial t} = \frac{\partial f}{\partial x} \frac{\partial \mathbf{v}}{\partial t} + \frac{\partial \rho}{\partial x} \left( \frac{\partial B}{\partial x} \right) - \frac{\partial \rho}{\partial x} \left( \frac{\partial B}{\partial x} \right) \left( \frac{\partial B}{\partial x} \right) \left( \frac{\partial B}{\partial x} \right),
\]

(2)

and the conservation of mass equation is

\[
\frac{\partial \rho}{\partial x} \left( \frac{\partial B}{\partial x} \right) = \frac{\partial \rho}{\partial t} \left( \frac{\partial B}{\partial t} \right) = \frac{\partial \rho}{\partial t} \left( \frac{\partial B}{\partial x} \right)
\]

(3)

where the bold variables represent time-varying complex variables or phasor quantities, \( \dot{m} \) is the mass flow rate (positive for flow from compressor to reservoir), \( x \) is the coordinate in the flow direction from
compressor to reservoir, $P$ is the pressure, $f_r$ is the Darcy friction factor, $r_h$ is the hydraulic radius, $A_g$ is the cross-sectional area of the gas perpendicular to the flow direction, $R$ is the gas constant on a mass basis, $T$ is the temperature, and $t$ is the time. The ideal-gas equation of state was used for the last term in Eq. (3). Equations (2) and (3) show phase relationships between flow and pressure. The first term on the right hand side of Eq. (2) represents the flow friction or flow resistance. The last term is important where $A_g$ changes rapidly. The second term on the right hand side of Eq. (2) was usually neglected in the early work with OPTRs since frequencies were low (<60 Hz) and mass flow rates were low. In those cases the pressure drop was in phase with the mass flow. However, at higher frequencies or higher flow rates the second term is no longer negligible in the pulse tube or in other tubes used for the complete system. This second term gives rise to a component of the pressure drop that is in phase with the acceleration of the mass, which for a sinusoidally varying mass flow leads the mass flow by 90°. It represents an inerterance effect and is brought about because of the inertia of the oscillating gas. It is analogous to an inductive effect in electrical systems. When combined with the resistive term, the pressure leads the mass flow by something less than 90°. As discussed earlier, this phase shifting at the warm end of the pulse tube is desirable to reduce the magnitude of the flow in the regenerator.

Equation (3) shows that any element that has a finite volume causes a change in mass flow toward the compressor that leads the pressure by 90°, which is opposite to the desired phase shift. Such a phase shift is known as a compliance effect and is analogous to a capacitance effect in electrical systems. The overall change in phase angle from one end to the other in any element is determined by solving both Eqs. (2) and (3). The beneficial phase shift caused by an inerterance effect within the pulse tube was first observed in 1996 by Godshalk et al. in an OPTR operating at 350 Hz. The inerterance effect can be enhanced further by using a long, narrow tube between the pulse tube and the reservoir, as shown in Fig. 6. The primary orifice in Fig. 6 can actually be removed and the entire flow impedance incorporated into the inerterance tube. In a pulse tube providing 19 W at 90 K, Marquardt and Radebaugh increased the phase shift further by using an inerterance tube with two diameters to provide a phase shift of 43°. The length of both tubes combined was 4.3 m, with the small-diameter tube next to the pulse tube. In much smaller pulse tubes the phase shift is so small that it is of little use, but in much larger pulse tubes the phase shift can be larger than optimum unless the geometry of the inerterance tube is adjusted to reduce the phase shift. In small pulse tubes both the secondary orifice and the inerterance tube need to be used together to give a desirable phase shift. The advantage of using only the inerterance tube is that there is no possibility of DC flow.

3.2. Efficiency of Pulse Tube Cryocoolers

3.2.1. Intrinsic OPTR efficiency

In the ideal OPTR the only loss is the irreversible expansion through the orifice. The irreversible entropy generation there is a result of lost work that otherwise could have been recovered and used to help with the compression. All other components are assumed to be perfect, and the working fluid is assumed to be an ideal gas. The $COP$ for this ideal OPTR is given by

$$COP_{ideal} = \frac{\dot{Q}_c}{W_0} = \frac{\langle P_d V_c \rangle}{\langle P_d V_h \rangle} = \frac{T_c}{T_h},$$

(4)

where $P_d$ is the dynamic pressure, $V$ is the volume flow rate, and the net refrigeration power is simply taken as the acoustic power at the cold end. The $\langle \rangle$ symbols indicate a time-averaged quantity. The acoustic power at the hot end of the regenerator is simply the PV power of the compressor. Because the regenerator is assumed to be perfect, the acoustic power varies along its length in accordance with the specific volume, which is proportional to temperature for an ideal gas. The maximum $COP$ from Eq. (4) is 1.0, but only when the cold temperature becomes equal to the hot temperature. A comparison of the $COP$ from Eq. (4) with the Carnot $COP$ from Eq. (1) shows that the only difference is the presence of the $T_c$ term in the denominator of Eq. (1). That term represented the work reversibly recovered at the low temperature and used to help in the compression. The Carnot efficiency of the ideal OPTR is given by
\[
\eta_{\text{ideal}} = \frac{\text{COP}_{\text{ideal}}}{\text{COP}_{\text{Carnot}}} = \frac{T_h - T_c}{T_h}.
\]

For \(T_h = 300 \text{ K}\) and \(T_c = 75 \text{ K}\), \(\eta_{\text{ideal}} = 0.75\). Since practical pulse tube refrigerators have efficiencies less than about 20\% of Carnot, the intrinsic loss is dominated by other practical losses when operating at this low temperature. However, for \(T_c = 250 \text{ K}\), \(\eta_{\text{ideal}} = 0.17\). In that case the lost power at the orifice is a much larger fraction of the total input power. Thus, the OPTR cannot compete with the vapor-compression refrigerator for near-ambient operation. It is useful only for much lower temperatures, especially cryogenic temperatures, unless the acoustic power flow at the warm end of the pulse tube is recovered.

### 3.2.2. Efficiencies of real pulse tube cryocoolers

Figure 7 shows a comparison of the efficiency of actual pulse tube refrigerators that have achieved high efficiencies. The efficiencies reported here refer to the input electrical power to the compressor. In a few cases the number 85 associated with a data point means that the efficiency was based on compressor PV power that was divided by 0.85 to obtain the electrical input power if an 85\% efficient compressor had been used. Such compressor efficiencies are typical of well-designed units. The majority of pulse tube refrigerators reported in the literature have not achieved efficiencies anywhere near these values. Careful attention to details in the design of these high efficiency refrigerators is required and expensive experimental optimization is often required even after the most careful computer modeling and optimizations are performed. In most cases these detailed designs remain proprietary information. Shown for comparison are data for recent, high efficiency Stirling, Gifford-McMahon, Brayton, and mixed-refrigerant Joule-Thomson refrigerators, all operating at temperatures near 80 K. This graph shows a general trend in all refrigerators for increased efficiency as the size increases. The graph also indicates that pulse tube refrigerators have equaled or exceeded the efficiency of the best Stirling refrigerators. As a result, pulse tube refrigerators have now become the most efficient cryocoolers for a given size. Efficiencies as high as 24\% of Carnot have now been achieved with pulse tube cryocoolers. As expected, the use of a valved compressor reduces the efficiency of the pulse tube refrigerator to about that of Gifford-McMahon refrigerators.

The minimum efficiencies required for SADA applications are also shown in Fig. 7, but they are much less than what has been achieved in many pulse tube cryocoolers. However, most actual Stirling cryocoolers for this application also have considerably higher efficiencies. The lower cost compressors for
the tactical military application are usually only about 65% efficient compared with the 85% typical of compressors for most space cryocoolers. With such compressors most of the Stirling-type pulse tube efficiencies would be reduced from that shown in Fig. 7. However, the data point at 40 W of input power at an efficiency of 12% of Carnot was a pulse tube cryocooler driven with a linear compressor intended for use with a Stirling cryocooler that provides at least 1.75 W at 67 K. The SADA requirement for the maximum input power is 94 W. With a pulse tube cold head the input power was 40 W when providing 1.75 W at 80 K, which is the same as the input power to the system when the original Stirling coldhead was in place\textsuperscript{29}. This pulse tube had a U-tube geometry that would not fit inside the envelope specified for the SADA applications.

The earliest pulse tube refrigerator to achieve these high efficiencies was the large NIST pulse tube in 1991 that achieved 31 W at 80 K with 602 W input PV power at 316 K (13% of Carnot assuming an 85\% efficient compressor). Within about a year the mini pulse tube refrigerator discussed earlier achieved about 8\% of Carnot\textsuperscript{21}. Most of the other pulse tube refrigerators with high efficiencies have been developed in the last few years for space applications with funding provided either by NASA or by the U.S. Air Force. The system with 17\% efficiency at 222 W of input power is the NIST oxygen liquefier\textsuperscript{18} mentioned previously that used the coaxial arrangement. Figure 8 shows this pulse tube cryocooler. Figure 9 shows one of the latest space-qualified pulse tube cryocoolers designed for cooling infrared sensors on the Integrated Multispectral Atmospheric Sounder (IMAS)\textsuperscript{30}. It provides 1.0 W cooling at 55 K and 2.5 W at 80 K with only 51 W of input power to the compressor\textsuperscript{31} to give an efficiency of 13.5\% of Carnot. It uses the latest technology in flexure-bearing compressors to reduce the size and mass of the compressor. Total mass of the compressor and pulse tube cold head is only 3.2 kg. It uses an inline arrangement for the regenerator and pulse tube.

One of the few disadvantages of pulse tube cryocoolers is the potential for gravitationally induced convective instabilities inside the pulse tube whenever the cold end of the pulse tube is raised above the warm end. The effect can be particularly pronounced in the off state, but the oscillating flow in the on state prevents the instability except for pulse tube diameters greater than about 10 mm. Very few measurements have been made of this instability in the on state. Two measurements in our laboratory, one with a 9-mm diameter pulse tube and the other with a 19-mm diameter pulse tube showed no orientation dependence for the first and about a 15 K temperature increase for the second. We expect that even for the 1.75 W SADA application the pulse tube diameter will be less than 10 mm, so this potential problem may not exist for these infrared applications.

---

**Figure 8.** Photograph of a coaxial pulse tube cryocooler for studies of liquefaction of oxygen on Mars.
4. COOLING INFRARED DETECTORS FOR TACTICAL APPLICATIONS

4.1. Requirements for the Standard Advanced Dewar Assembly (SADA)

Table 2 lists the performance and geometry requirements for cryocoolers used with second-generation thermal imaging systems. As discussed in the previous section, the efficiency of many current pulse tube refrigerators greatly exceeds the SADA requirements. The four remaining challenges in the use of pulse tubes for this application are (1) the geometry of the cold finger, (2) the operation at any orientation, (3) the cool-down time, and (4) the additional mass and volume due to the reservoir. The many other SADA requirements not listed in Table 2 should affect the pulse tube cryocooler no differently than they affect the Stirling cryocooler. In fact, the vibration levels with the pulse tube cryocooler should be significantly less

<table>
<thead>
<tr>
<th>Requirement</th>
<th>System</th>
<th>0.15 W</th>
<th>0.35 W</th>
<th>1.0 W</th>
<th>1.75 W</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cooling Capacity @ 23 °C Ambient</td>
<td></td>
<td>0.15 W</td>
<td>0.35 W</td>
<td>1.0 W</td>
<td>1.75 W</td>
</tr>
<tr>
<td>@ 77 K</td>
<td></td>
<td>@ 77 K</td>
<td>@ 80 K</td>
<td>@ 77 K</td>
<td>@ 67 K</td>
</tr>
<tr>
<td>Cool-down Time to 80 K</td>
<td></td>
<td>2.5 min</td>
<td>10 min</td>
<td>13 min</td>
<td>6.5 min</td>
</tr>
<tr>
<td>w/125 J</td>
<td></td>
<td>w/250 J</td>
<td>w/1440 J</td>
<td>w/1200 J</td>
<td></td>
</tr>
<tr>
<td>Input Power (max)</td>
<td></td>
<td>17 W</td>
<td>35 W</td>
<td>60 W</td>
<td>94 W</td>
</tr>
<tr>
<td>Mass (max)</td>
<td></td>
<td>0.45 kg</td>
<td>1.14 kg</td>
<td>1.91 kg</td>
<td>3.41 kg</td>
</tr>
<tr>
<td>Reliability, Min MTTF</td>
<td></td>
<td>4000 hr</td>
<td>4000 hr</td>
<td>4000 hr</td>
<td>4000 hr</td>
</tr>
<tr>
<td>Cold Finger Diameter</td>
<td></td>
<td>6.6 mm</td>
<td>5.0 mm</td>
<td>13.7 mm</td>
<td>13.7 mm</td>
</tr>
<tr>
<td>Cold Finger Length</td>
<td></td>
<td>57 mm</td>
<td>62 mm</td>
<td>59 mm</td>
<td>59 mm</td>
</tr>
</tbody>
</table>
than those with the Stirling cryocooler because of the lack of moving parts in the cold head. With the SADAs the sleeve for the cryocooler cold finger is built into the dewar. For a Stirling cold head the displacer fits inside this sleeve as shown in Fig. 10a. Figure 10b shows how a coaxial pulse tube would fit into the same sleeve.

4.2. Adapting Pulse Tube Cryocoolers to SADA Requirements

The optimum cross-sectional area of the regenerator in any cryocooler will be proportional to the refrigeration power to a first approximation, and the length as well as the mesh size of the screen regenerator will not change. This scaling law assumes all losses are proportional to the cross-sectional area for a fixed flow velocity. This assumption is reasonably good until radiation loss begins to dominate. The cross-sectional area of the pulse tube can also be scaled with the refrigeration power until at some small size the thermal penetration depth in the gas is no longer small compared with the tube radius. At 60 Hz the thermal penetration depth in helium at 300 K is 0.20 mm. We have found that for a pulse tube radius less than about 2.0 mm performance deteriorates quickly because of excessive heat transfer between the gas and the wall. With this size of system the radiation loss also begins to become quite important since it scales with the surface area instead of the cross-sectional area. Most published data on the pulse tube cryocoolers with high efficiencies do not give dimensional data on the regenerators and pulse tubes. However, data from two NIST pulse tube cryocoolers with high efficiencies is available and has been published. The first is the pulse tube oxygen liquefier that produced 15 W at 80 K using a coaxial geometry, and the second is an inline pulse tube cryocooler that produced 2.5 W of cooling at 80 K using the linear compressor from the tactical Stirling cryocooler intended to produce 1.75 W at 80 K. More recent unpublished data with the same compressor and regenerator #4, but with a larger pulse tube (5.33 mm ID x 45.7 mm long) yielded 3.4 W of cooling at 80 K. The compressor PV power at this high refrigeration power was about 100 W and is greater than the SADA specifications. Nevertheless, this data along with the data from the oxygen liquefier can be used to determine the coefficients in the scaling of cross-sectional areas of the regenerators and pulse tubes.

For the regenerator the ratio of inside cross-sectional area to net refrigeration power at 80 K is $A_{rq} = 0.32 \text{ cm}^2/\text{W}$ for the 15 W system and $0.28 \text{ cm}^2/\text{W}$ for the 2.5 W system. To be conservative the higher value will be used for scaling to smaller sizes. The regenerator length was 50 mm for the larger one and 40 mm for the smaller one. Performance of the larger system with a 40-mm long regenerator was almost the same. For the pulse tube the ratios are $A_{pq} = 0.084 \text{ cm}^2/\text{W}$ for the large system and $0.057 \text{ cm}^2/\text{W}$ for the small system. Again, the larger value will be used to be conservative regarding a fit inside the SADA sleeve. The pulse tube lengths were 70 mm for the larger system and 46 mm for the smaller system. Ideally the pulse tube length in a coaxial system should be somewhat longer than the regenerator to match the temperature profiles closely. An extra 10-mm length in the pulse tube compared with the regenerator can be located at the warm end and outside the SADA sleeve. Sleeve lengths for the four SADA systems are in the range of 55 to 60 mm with about 10 mm at the warm end with no temperature gradient.

![Figure 10. Integration of cryocooler with standard dewar assembly.](image-url)
With the specific areas given above for the regenerator and the pulse tube along with a pulse tube wall thickness of 0.15 mm and a regenerator wall thickness of 0.25 mm, we obtain the diameters as given in Table 3. For the two smaller systems the pulse tube diameters given in the parentheses indicate the diameters when using the scaling relationship. The 3.0 mm diameter is given next as the minimum allowed diameter to prevent excessive heat transfer with the wall. Regenerator areas were increased by 20% for the 0.15 W system and 10% for the 0.35 W system to allow for degradation of net refrigeration at small sizes. The regenerator areas in parentheses are the values before the 205 and 10% were added. As this table shows, the cold fingers predicted from simple scaling laws applied to previously tested pulse tube cryocoolers have diameters smaller than the SADA requirements, except for the 0.35 W cooler. It may still be possible to meet the requirements for the 0.35 W cooler, but more careful optimization and experimentation would be required. The scaling law used here was for 80 K, but the 1.75 W cooler must provide that cooling at 67 K. Thus more PV power will be required and may increase the diameter of the cold finger some, but it should still be possible to keep it less than 13.7 mm. Because the pulse tube diameter even for the largest cooler is significantly less than 10 mm in diameter, there should be no orientation dependence during operation of any of the four pulse tube coolers. Cool-down times for pulse tube cryocoolers should be quite similar to that of an equivalent Stirling cryocooler because the size and mass of the cold finger is about the same for the two systems. Table 3 also lists the approximate volumes needed for the reservoirs of the various pulse tube coolers. This volume is taken as 50 times the pulse tube volume. Often this reservoir volume is placed internal to the compressor and requires no additional volume and a very small additional mass. That procedure requires that the compressor be designed specifically to accommodate the reservoir. It could also be place external to the compressor as an annular sleeve. An external reservoir could also be made part of the warm end of the cold head, including the use of fins for heat rejection. This extra volume is not significantly more than that required for the pneumatic drive of the displacer in Stirling cryocoolers. Details of the envelope specification for the warm end of the cold head would need to be examined in detail to be able to best incorporate the reservoir volume.

5. CONCLUSIONS

Many advances in pulse tube refrigerators have occurred in the last few years, particularly in methods to increase the efficiency. Several high efficiency pulse tube cryocoolers have been developed for space applications, mostly for cooling infrared detectors. Efficiencies in the range of 10 to 24% of Carnot at 80 K have been achieved in small pulse tube cryocoolers with input powers of 10 to 100 W. Such efficiencies are much higher than the 2.5 to 6.4% efficiencies required of the cryocoolers for the second-generation thermal imaging systems in the Standard Advanced Dewar Assembly (SADA) specifications. With such high efficiencies it should then be possible to make pulse tube refrigerators with coaxial geometries for the pulse tube and regenerator that have small enough diameters and lengths for the cold finger to fit SADA systems. Estimated diameters were determined from scaling laws applied to somewhat larger systems that were very efficient and well characterized. The 1.0 W and 1.75 W coolers should be relatively easy to retrofit. The 0.35 W cooler will require careful optimization and experimentation to develop a pulse tube cooler with such a small diameter cold finger. The 0.15 W cooler with maximum input power of 17 W is

Table 3. Estimated geometry of cold finger for coaxial pulse tube cryocooler.

<table>
<thead>
<tr>
<th>Geometry</th>
<th>System 0.15 W</th>
<th>System 0.35 W</th>
<th>System 1.0 W</th>
<th>System 1.75 W</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pulse Tube Area (cm²)</td>
<td>0.0126</td>
<td>0.0294</td>
<td>0.084</td>
<td>0.147</td>
</tr>
<tr>
<td>Pulse Tube ID (mm)</td>
<td>(1.27) 3.0</td>
<td>(1.93) 3.0</td>
<td>3.27</td>
<td>4.33</td>
</tr>
<tr>
<td>Pulse Tube OD (mm)</td>
<td>3.3</td>
<td>3.3</td>
<td>3.57</td>
<td>4.63</td>
</tr>
<tr>
<td>Regenerator Area (cm²)</td>
<td>(0.048) 0.058</td>
<td>(0.112) 0.123</td>
<td>0.32</td>
<td>0.56</td>
</tr>
<tr>
<td>Regenerator ID (mm)</td>
<td>4.27</td>
<td>5.15</td>
<td>7.31</td>
<td>9.63</td>
</tr>
<tr>
<td>Regenerator OD (mm)</td>
<td>4.77</td>
<td>5.65</td>
<td>7.81</td>
<td>10.03</td>
</tr>
<tr>
<td>Reservoir Volume (cm³)</td>
<td>(3.8) 21</td>
<td>(3.8) 21</td>
<td>25</td>
<td>44</td>
</tr>
<tr>
<td>Required Cold Finger OD (mm)</td>
<td>6.6</td>
<td>5.0</td>
<td>13.7</td>
<td>13.7</td>
</tr>
<tr>
<td>Required Cold Finger Length (mm)</td>
<td>56.8</td>
<td>61.6</td>
<td>58.5</td>
<td>58.5</td>
</tr>
</tbody>
</table>
near the lower limit of size for efficient pulse tube cryocoolers. Fortunately, the cold finger diameter is not particularly small for this cooler, but the challenge on this system is to maintain high efficiency with such a small compressor. The small pulse tubes in these systems should not experience any orientation dependence. Several possible locations for the reservoir volume were discussed.

ACKNOWLEDGEMENTS

I appreciate the valuable discussions with W. E. Salazar and R. M. Rawlings regarding the requirements and characteristics of linear Stirling cryocoolers in the cooling of infrared detectors. I also wish to thank R. M. Rawlings for the loan of two linear Stirling compressors for our use in the development of small pulse tube cryocoolers.

REFERENCES