Ultrahigh vacuum practice

G. F. Weston, MSc, FInstP
Philips Research Laboratories

Butterworths
London · Boston · Durban · Singapore · Sydney · Toronto · Wellington

First published 1985

British Library Cataloguing in Publication Data

Weston, G.F.
Ultrahigh vacuum practice.
1. Vacuum technology
I. Title
621.5'5 TJ940

ISBN 0-408-01485-7

Library of Congress Cataloging in Publication Data

Weston, G. F. (George Frederick)
Ultrahigh vacuum practice.

Includes bibliographies and index.
1. Vacuum technology. I. Title.
TJ940.W47 1985 621.5'5 84-23885
ISBN 0-408-01485-7

Filmset by Mid-County Press, London SW15
Printed and bound in Great Britain at the University Press, Cambridge

Preface

In the 1950s vacuum science and technology took a great leap forward. For the first time, pressures below 10^{-5} Pa (10^{-7} torr) could be attained and measured and the term 'ultrahigh' vacuum was coined to cover this low-pressure regime. The initial studies created considerable interest with the result there was a huge upsurge of research and development work in the field. By the 1960s pressures down to 10^{-11} Pa were being reached. Since most of the applications did not require pressures below 10^{-8} Pa, there seemed little need to push the technology any further and for the last twenty years there has been a period of consolidation. Ultrahigh vacuum components based on the designs of the 1950s and 1960s have been developed and engineered into reliable commercial products, so that ultrahigh vacuum can be exploited as a tool to provide the necessary environment for many research, development and production processes.

In the 1960s a number of textbooks were published on vacuum technology. Most of them covered the entire vacuum pressure range with limited information on ultrahigh vacuum. The few devoted to ultrahigh vacuum alone, concentrated mainly on the physical principles involved. Since the 1960s there have been hardly any books published on the subject, so that textbook information, particularly on hardware, now tends to be rather dated. In an attempt to fill this gap the author wrote a series of articles for *Vacuum* in their *Educational Series* at the beginning of the 1980s, aimed at giving practical information on materials and components such as pumps, gauges, valves, etc. for ultrahigh vacuum applications. The articles were well received, judging by the request for reprints, and it was suggested by his colleagues that the author should gather this information together in a textbook which would then be more readily available to the non-specialist user, who would not be familiar with conference reports and journals on the subject. Thus this book was born.

Ultrahigh Vacuum Practice is written mainly for the user or potential user of ultrahigh vacuum equipment, who is not a vacuum expert but is acquainted with general vacuum practice. However, in bringing the information together in one place with extensive references, it is hoped that even the expert vacuum engineer will find it a useful book to have on his shelf. It is basically a practical book, dealing with components suitable for ultrahigh vacuum applications, their theory of operation, their assembly and use and their performance and

calibration. In order to obtain maximum performance from vacuum equipment, the users must often have a reasonable understanding of the physical principles involved, so that where necessary background theory has been included. Thus Chapter 1 covers the fundamental principles of vacuum such as adsorption and desorption of gases from surfaces, diffusion, pumping mechanism and gas flow. Chapter 2 discusses materials, their preparation and joining methods necessary for work in this field. In Chapter 3 the various pumps available are detailed with performance and calibration techniques. Chapters 4 and 5 deal with total pressure and partial pressure gauges respectively, whilst Chapter 6 gives details of valves, seals and other vacuum line components. The design of systems using the components described in the previous chapters is discussed in Chapter 7, including also the procedures for making best use of the equipment. Chapter 8 discusses the problems of leak detection and describes suitable leak detectors.

The author gratefully acknowledges the cooperation and encouragement of his colleagues in the research laboratories. He is especially indebted to Dr L. G. Pittaway who read the manuscript and made many useful suggestions and corrected a number of errors.

G.F.W.
1985

Chapter 1

Fundamentals of vacuum science and technology

1.1 Properties of vacuum

Whenever the density of gas in a given volume is reduced below that corresponding to the density of atmospheric gas at ground level, a vacuum is obtained. The greater the reduction of the gas density, the better the vacuum. This reduction of density gives rise to a number of interesting properties which can be exploited in a variety of applications. For example, the chemical activity of atmospheric gas such as oxidation of metals is greatly reduced. Thus, under the protective environment of vacuum, reactive materials can be stored and their special properties utilized. If the density is low enough, surfaces can be kept clean for several hours without even a monolayer of gas depositing on them. Thus, atomically clean surfaces can be investigated and the effects of gas layers determined. The relative absence of molecules in a vacuum allows very small projectiles to pass unimpeded across large distances. This is of particular interest when dealing with charged particles such as electrons, ions and protons whose precise path in the absence of collisions can be controlled by electric and/or magnetic fields. Physical processes such as the passage of heat, of sound and indeed of the gas itself, which take place by interaction between the gas particles under atmospheric conditions, are greatly modified when the gas density is reduced, to the extent that interaction of the gas molecules is no longer the predominant mechanism of transport.

The various properties and effects mentioned above obviously depend on the degree of vacuum achieved. Thus the density of the gas remaining in the space serves directly as a measure of the degree of vacuum. However, since historically following on the work of Boyle, the density of a gas was known to be directly proportional to the gas pressure, in practice the gas pressure has become the accepted measure of the degree of vacuum.

Modern vacuum techniques can now produce a vacuum in which the pressure has been reduced by a factor of 10^{15} compared with atmospheric pressure. This is a very large range of vacuum and it is convenient to subdivide it. *Figure 1.1* is a chart showing the generally accepted terms for the subdivisions, and the corresponding pressures expressed in Pascals*. Also

* The Pascal is the special name for the Newton per square metre and is the SI unit for pressure. Previously mm of mercury or torr have been commonly used and often millibars are used instead of the SI units. A conversion table relating the various pressure units is given in Appendix 3

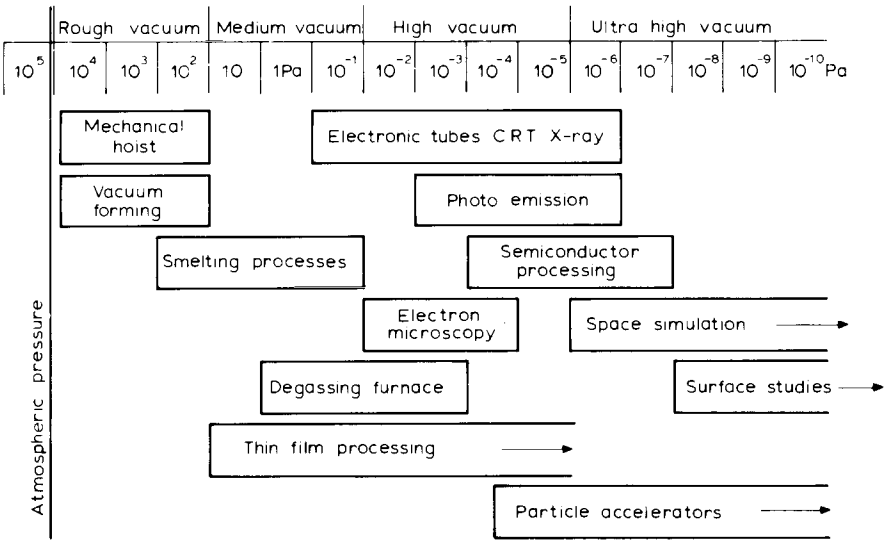


Figure 1.1 Vacuum scale and applications

included in *Figure 1.1* are some examples of practical applications which utilize the special properties of a vacuum. The application to mechanical hoists arises not because of any special property of a vacuum, but merely because the pressure difference which exists across the walls of a vacuum container can give rise to large forces which can be harnessed in mechanical devices by proper design.

The exploitation of the vacuum properties outlined above thus requires provision of the correct vacuum environment and this in turn requires employing the correct equipment in a satisfactorily designed system. To design and derive the optimum performance from a vacuum system, the engineer not only needs to know the performance of the equipment but also the factors which will affect it. It is not enough to know that a pump has a speed of say $10^{-1}\text{m}^3\text{s}^{-1}$ and can reach an ultimate pressure of 10^{-6} Pa. In a badly designed system the performance can fall below the optimum by more than an order of magnitude. Indeed, he needs to understand the fundamental principles of vacuum technology to get the ultimate from his equipment. This is particularly true in the ultrahigh vacuum region (pressures below 10^{-6} Pa) where, for example, the number of gas molecules adsorbed on the surface of the vacuum chamber can far exceed the number in the space itself. This chapter deals with these fundamental principles. The subject is covered fairly briefly as a background to the chapters that follow, and for a more detailed presentation the reader is referred to the excellent book by Redhead *et al.*[1].

1.2 The kinetic theory model of a vacuum

A vacuum environment is a gas at reduced pressure and thus any theory of vacuum must account for the usual laws which have been shown by experiment to hold true for gases. These laws, i.e. Boyle's Law, Charles' Law

and Avogadro's Hypothesis, can all be combined into one statement called the equation of state, namely

$$PV = NkT \tag{1.1}$$

where P is the gas pressure, V the volume, T the absolute temperature and N the total number of molecules. k is Boltzmann's constant, which is applicable universally to all gases and is equal to 1.38×10^{-16} ergs per K: for the vacuum engineer it is better expressed as $k = 1.38 \times 10^{-23}$ Pa $m^2 K^{-1}$.

In addition to the above experimentally observed relationship, the theory should account for the pressure dependence of the various properties which have been mentioned in Section 1.1. The kinetic theory of gases has been most successful in this respect. The theory assumes that the molecules of a gas can be represented by hard spheres in continual random motion. Inter-molecular collisions, as well as the collisions of molecules with the walls of the vacuum vessel, are assumed to take place elastically. Starting from these basic assumptions, a number of relationships and criteria have been deduced which have shown that the kinetic theory can account for the properties of gases. Furthermore it can predict the correct dependence on pressure of such physical properties of gases as viscosity and thermal conductivity.

In the following sections those relationships which are particularly relevant to a discussion of high and ultrahigh vacuum techniques will be outlined, whilst the detailed derivation of these important equations can be found in Appendix 1. For further information the reader is referred to books on the subject of the kinetic theory of gases, such as that by Present[2].

1.2.1 Pressure

By considering the rate of change of momentum of the randomly moving molecules striking a surface, the expression for the pressure of the gas is found to be

$$P = \tfrac{1}{3} m n v_{\text{rms}}^2 \tag{1.2}$$

where m is the mass of a molecule, n the number of molecules in unit volume and v_{rms} is the root mean square of all the possible molecular velocities. The derivation of Equation (1.2) is given in Appendix 1, Section A1.1.

1.2.2 Energy and velocity of the molecules

Dividing Equation (1.1) by V yields

$$P = nkT \tag{1.3}$$

and thus by comparing Equations (1.2) and (1.3), it is seen that in order for the kinetic theory to agree with the experimentally observed result

$$v_{\text{rms}} = \sqrt{\frac{3kT}{m}} \tag{1.4}$$

Since the kinetic energy of a particle of mass m, moving with velocity v_{rms}, is $\frac{1}{2} m v_{\text{rms}}^2$, it follows, from Equation (1.4), that the mean kinetic energy of a

molecule is 3 kT/2. The kinetic theory thus associates the mean kinetic energy of a molecule with the absolute temperature.

The continual collisions between the molecules result in a definite distribution of molecular velocities for a body of gas in a steady state (see Appendix 1). Under these conditions the number of molecules per unit volume, dn, having velocity components lying between v and $v+dv$, within a solid angle $d\omega$ is given by

$$dn = nf(\bar{v})\,dv d\omega = nf(\bar{v})\,dv \sin\theta\, d\theta d\phi \tag{1.5}$$

where n is the total number of molecules per unit volume, $\bar{v}$ is the vector velocity giving both speed and direction and $f(\bar{v})$ is the distribution function.

From statistical–mechanical considerations, Maxwell showed that the distribution function could be expressed in polar coordinates as

$$f(\bar{v}) = \left(\frac{m}{2\pi kT}\right)^{3/2} \exp(-mv^2/2kT)v^2 \sin\theta$$

Of more interest is the velocity distribution in any direction $f(v)$ which is obtained by integrating over all directions

$$f(v) = \left(\frac{m}{2\pi kT}\right)^{3/2} \exp(-mv^2/2kT)v^2 \sin\theta \tag{1.6}$$

from which the mean velocity

$$v_m = 2\sqrt{\frac{2kT}{\pi m}} \tag{1.7}$$

(see Appendix 1, Section A1.2).

1.2.3 Rate of incidence of molecules on a surface

It is frequently useful to know the rate v at which molecules strike unit area of a surface placed within the vacuum. This has been derived from kinetic theory in Appendix 1, Section A1.3 as

$$v = \tfrac{1}{4} n v_m \tag{1.8}$$

Introducing the appropriate expressions for v_m and n, Equation (1.8) can be expressed in the alternative form

$$v = \frac{P}{\sqrt{2\pi mkT}} \tag{1.9}$$

1.2.4 Mean free path

As a result of the random motion of the molecules, collisions will occur between them and it will be possible to calculate the mean distance travelled by the molecules between collisions. This average distance, called the mean free path, is seen from Appendix 1, Section 1.4 to be given for collisions between similar molecules in a gas under steady state conditions, by the expression

$$\lambda = \frac{1}{n\pi d^2\sqrt{2}} \tag{1.10}$$

where d is the diameter of the sphere representing the molecule.

As the pressure or density of the gas is reduced, the mean free path increases, so that at some pressure the mean free path will be large compared to the dimensions of the vacuum vessel. Under these conditions, collisions between molecules and the walls of the vessel will be more frequent than intermolecular collisions. The passage of gas through the vacuum system is then said to occur under conditions of free molecular flow. *Table 1.1* contains values of the mean free path for nitrogen at different pressures. Also included in this table are values of ν, the rate of incidence of molecules on unit area and the time taken for a monolayer of gas to be deposited on a surface.

TABLE 1.1. Values of the parameters n, ν, λ and τ at different pressures

Pressure (Pa)	10^5 (*atmospheric*)	10^{-4}	10^{-6}	10^{-8}	10^{-10}
(Torr)	750	7.5×10^{-7}	$7,5 \times 10^{-9}$	7.5×10^{-11}	7.5×10^{-1}
n	2.7×10^{19}	2.7×10^{10}	2.7×10^{8}	2.7×10^{6}	2.7×10^{4}
ν	2.8×10^{23}	2.8×10^{14}	2.8×10^{12}	2.8×10^{10}	2.8×10^{8}
λ	6×10^{-6}	6×10^{3}	6×10^{5} ($\sim$4 miles)	6×10^{7}	6×10^{9}
τ	3×10^{-19} s	3 s	5 min	$8\frac{1}{2}$ h	35 days

n = number of molecules per cm^3 at 0°C
ν = number of molecules incident on cm^2 at 0°C s^{-1}
λ = mean free path in cm at 0°C
τ = time for a monolayer to form assuming a sticking probability of unity

1.2.5 Thermal transpiration

Under conditions of free molecular flow, the criterion for equilibrium of the gas in two large chambers connected by an orifice of dimensions much less than the mean free path, is simply that the number of molecules passing in each direction through the orifice in unit time should be equal. This number is ν, given by Equation (1.9), multiplied by the area of the orifice. Hence, when the two chambers are held at different temperatures T_1 and T_2, equilibrium will be established only when

$$\frac{P_1}{\sqrt{T_1}} = \frac{P_2}{\sqrt{T_2}} \tag{1.11}$$

i.e. for equilibrium under these conditions it is necessary that a pressure difference also exists.

1.3 Flow of gas through vacuum systems

1.3.1 Mass flow

The process of evacuation involves the removal of a mass of gas from the vacuum vessel. The rate of removal, i.e. the mass flow, determines the rate at which the pressure falls.

If the vessel initially contains N molecules of mass m, the rate of change of mass is given by

$$\frac{dM}{dt} = \frac{d(Nm)}{dt} \tag{1.12}$$

Substituting for N from Equation (1.1)

$$\frac{dM}{dt} = \frac{d\left(\frac{mPV}{kT}\right)}{dt} \tag{1.13}$$

In practice, the type of gas and the temperature can be considered to be constant during evacuation, so that

$$\frac{dM}{dt} = \frac{m}{kT}\frac{d(PV)}{dt} \tag{1.14}$$

Both the pressure and volume are easily measured quantities in vacuum vessels so that it is convenient to define a gas flow rate Q as

$$Q = \frac{d(PV)}{dt} \tag{1.15}$$

Q can then be measured in practical cases and is related to the mass flow rate by

$$\frac{dM}{dt} = \frac{m}{kT}Q \tag{1.16}$$

provided always that m and T are constant. Q has the units of Pa $m^3 s^{-1}$ and is often referred to as gas throughput.

1.3.2 Conductance

Under constant temperature conditions, a flow of gas through a hole or pipe will always occur when a pressure difference exists.

Under conditions of free molecular flow, the gas flow rate Q through the duct is observed to be directly proportional to the pressure difference, i.e.

$$Q = C(P_1 - P_2) \tag{1.17}$$

where C is the constant of proportionality, which depends on the physical dimensions of the duct and is known as the conductance. Equation (1.17) is therefore a definition of conductance.

It follows from this definition that when two conductances C_1 and C_2 are connected together, the total conductance C is given by

$$C = C_1 + C_2 \quad \text{for parallel connection} \tag{1.18}$$

$$\frac{1}{C} = \frac{1}{C_1} + \frac{1}{C_2} \quad \text{for series connection} \tag{1.19}$$

1.3.3 Conductance of an orifice

Suppose as in Section 1.2.5, that two large chambers are connected by a small orifice, the orifice having no thickness and having a diameter much smaller than the mean free paths of the molecules. If the chambers are chosen to be large so that the flow of gas through the hole makes an insignificant difference to the distribution of molecules in the chamber, then the pressure throughout each individual chamber can be assumed to be uniform.

Assuming both chambers to be at the same temperature but the pressure in one to be raised to a higher pressure, then under steady flow conditions the nett number of molecules flowing from the high pressure to the low pressure chamber in unit time is given by

$$\frac{(P_1 - P_2)A}{\sqrt{2\pi mkT}}$$

from Equation (1.9), where A is the area of the orifice. The mass rate of flow is therefore

$$\frac{dM}{dt} = \frac{(P_1 - P_2)Am}{\sqrt{2\pi mkT}} \tag{1.20}$$

Using Equation (1.16), the gas flow Q is thus

$$Q = \sqrt{\frac{kT}{2\pi m}}(P_1 - P_2)A \tag{1.21}$$

Comparing Equation (1.21) with (1.17) the conductance of the orifice is therefore

$$C_0 = A\left(\frac{k}{2\pi}\right)^{1/2}\left(\frac{T}{m}\right)^{1/2} \tag{1.22}$$

Thus, the conductance of an orifice is a function of temperature, the type of gas and the area of the orifice.

1.3.4 Conductance of tubes

The flow of gas through a tube under conditions of free molecular flow implies that the molecules will collide mainly with the walls of the tube as they pass through and hardly ever with one another. It is usually assumed that the walls of the tube are very rough in terms of molecular dimensions so that although the collisions of the molecules with the walls are elastic, there is no simple reflection of the molecules. Instead, the molecules leaving a small area of the surface are assumed to be diffusely scattered, so that the distribution of molecules travelling in a direction making an angle θ to the normal to the surface varies as the cosine of the angle θ.

These assumptions lead (see Present[2]) to an expression for the flow across any section of an infinitely long tube of radius r given by

$$Q = \tfrac{4}{3}r^3\left(\frac{2\pi kT}{m}\right)^{1/2}\frac{dP}{dx} \tag{1.23}$$

If the tube is not infinite, but still sufficiently long that the influence of the open ends is negligible, the flow rate across any section can still be assumed constant and hence dP/dx can be replaced by $(P_1 - P_2)/L$ where L is the length of the tube. Thus the conductance of a long tube is given by

$$C_{LT} = \frac{4}{3}\frac{r^3}{L}(2\pi k)^{1/2}\left(\frac{T}{m}\right)^{1/2} \tag{1.24}$$

For short tubes in which the area of the open ends is appreciable compared with the wall area, Dushman[3] has argued that an approximate value for their conductance can be obtained by considering the combined conductance of the tube and an orifice.

Using Equation (1.19) for conductances connected in series, Equations (1.22) and (1.24) can be combined to give an overall conductance of the short tube given by

$$C_{ST} = \frac{C_0}{1 + \frac{3}{8}\frac{L}{r}} \tag{1.25}$$

Clausing[4] has calculated the conductance of short tubes directly and *Figure 1.2* contains a chart based on Clausing's results which enables the conductance of short tubes to be read off. More accurate calculations for circular tubes by Cole[5] using modern computer techniques confirm that the early data of Clausing were completely satisfactory for most practical purposes.

When combining a number of conductances in order to obtain an overall conductance figure, it should be realized that the calculations are based on the assumption that the ends of the tubes are open to larger areas, i.e. not restricted, and that a uniform pressure gradient occurs along the tube. In

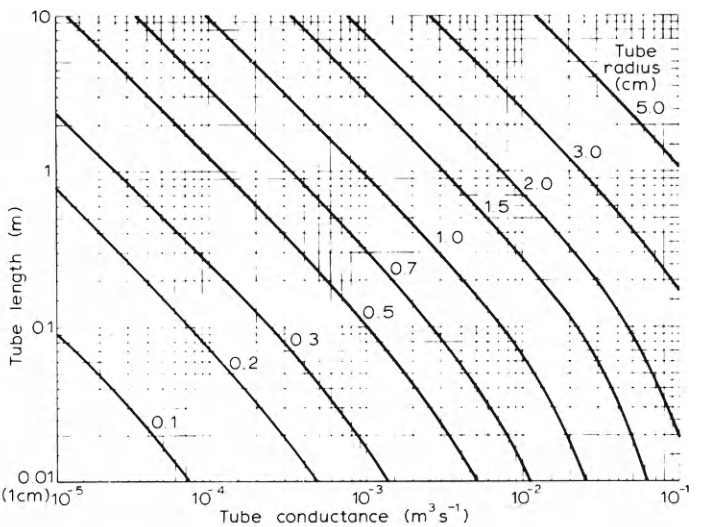


Figure 1.2 Conductance of cylindrical tubes under molecular flow conditions according to Clausing[4]

practice these conditions will not apply, particularly if the diameter of the tube varies, and inaccuracies will occur.

In cases where an accurate knowledge of the conductance of complex shapes is required, it has been shown that methods of computation can be applied which trace the path of molecules through the system and determine for a large number of molecules the probability of transmission. Examples of this method, called Monte Carlo, are to be found in papers by Davis[6] and Chubb[7].

1.4 Pumping of vacuum systems

1.4.1 Pump speed

The purpose of a pump is to provide a means for removing gas. Many practical pumps operate continuously so that if a constant flow of gas Q is introduced into the pump then a steady state is set up with an associated constant pressure within the pump of P.

If we define a quantity S_0 by

$$S_0 = Q/P \tag{1.26}$$

then S_0 is clearly a measure of the pumping ability of the pump. S_0 is called the pump speed and has the useful property that for many pumps it remains fairly constant over the working pressure range of the pump. Comparing Equation (1.26) with Equation (1.17), pump speed is seen to have the same dimensions as conductance and in SI units is measured in m^3s^{-1}. If it is desired that the pump should be connected to a vacuum chamber by means of a conductance C, then the pump speed available at the chamber can be deduced as follows.

The flow of gas through the conductance produces a pressure drop given by Equation (1.17) as

$$Q = C(P_1 - P_2) \tag{1.27}$$

where P_1 is the pressure in the system and P_2 that at the pump. The flow of gas into the pump is given by Equation (1.26) as

$$Q = S_0 P_2 \tag{1.28}$$

The pump speed in the chamber is defined by

$$S = Q/P_1 \tag{1.29}$$

Eliminating Q, P_1 and P_2 from these equations gives

$$S = \frac{C}{1 + \dfrac{C}{S_0}} \tag{1.30}$$

It follows that for any fixed conductance C, whatever value of pump speed S_0 is chosen, the maximum pump speed available in the chamber does not exceed C m^3s^{-1}. Only when the pump is fitted in the chamber, i.e. $C = \infty$, can all the pump speed by utilized and even when the size of the conductance C equals the pump speed S_0, the speed available in the chamber is only half, i.e. $S_0/2$.

1.4.2 Rate of evacuation

From the principle of conservation of matter, the rate of change of mass of gas in the gas phase in a vacuum vessel is equal to the difference between the mass of gas entering and leaving the gas phase in the vessel in unit time. If the temperature is assumed to be constant and the type of gas is also constant, the rate of mass change is proportional to the flow rate (Equation (1.15)) expressed in units Pa $m^3 s^{-1}$ and the equivalent methematical statement for conservation becomes

$$\frac{d(PV)}{dt} = Q_i - Q_o \tag{1.31}$$

where the subscripts i and o refer to the flow rate into and out of the chamber, respectively. If the pumping speed available in the chamber is S, the rate of outflow is given by

$$Q_o = SP \tag{1.32}$$

For a vacuum system of constant volume, Equation (1.31) then becomes

$$V\frac{dP}{dt} = Q_i - SP \tag{1.33}$$

If it is assumed that Q_i is constant, then Equation (1.33) may be integrated directly to give

$$P = \frac{Q_i}{S} - \left(\frac{Q_i}{S} - P_0\right)\exp\left(-\frac{S}{V}t\right) \tag{1.34}$$

where P_0 is the initial pressure. During pump down the pressure therefore falls exponentially with a time constant V/S and at $t = \infty$ reaches an ultimate pressure given by

$$P_{ult} = \frac{Q_i}{S} \tag{1.35}$$

1.4.3 Pump speed, gas influx and the degree of vacuum

Equation (1.34) shows clearly that the pump speed and the rate of gas influx are intimately related to both the speed of evacuation and the ultimate pressure achieved. This relationship is the essence of vacuum technique and has far reaching implications. In order to demonstrate its meaning in more practical terms, consider a vacuum vessel consisting of a hollow cube one metre in height. Suppose the total rate of flow of gas entering the system from all sources is Q_i.

In order to remove the gas it is necessary to provide a pump within the system. If we consider a surface within the vessel then molecules strike this surface at a rate given by Equation (1.9) as

$$v = \frac{P}{\sqrt{2\pi mkT}}$$

If it were possible to make all the molecules which strike the surface stick to it permanently, this would provide a means of pumping and, furthermore, it

would provide the fastest possible pump because the molecules cannot reach the surface at a faster rate than v.

The gas flow to the surface per unit area Q_A is obtained from Equation (1.16) as

$$Q_A = \frac{kT}{m}\frac{dM}{dt} = vkT$$

and substituting for v

$$Q_A = P\sqrt{\frac{kT}{2\pi m}} \tag{1.36}$$

Hence, from the definition of pump speed, the pump speed per unit area of the hypothetical surface to which all gas sticks is

$$S_A = \frac{Q_A}{P} = \sqrt{\frac{kT}{2\pi m}} \tag{1.37}$$

Inserting the numerical values corresponding to nitrogen gas at room temperature (cf. *Table 1.1*), the maximum possible pumping speed per unit area is 100 $m^3 s^{-1}$.

In the cubic metre vessel taken as the practical example it would be convenient to set aside an area of 10^{-2} m^2 on one face of the cube and consider this area to act as the perfect pump. The maximum pumping speed available in the vessel would then be 1 $m^3 s^{-1}$. It is now possible to calculate the allowable magnitude of the total gas influx Q_i for any value of the desired ultimate pressure in the vessel using Equation (1.35). These data are shown in *Table 1.2* for the vacuum vessel considered in the practical example.

TABLE 1.2. Ultimate pressure as a function of gas load for the vacuum vessel used in the example

Q_i (Pa $m^3 s^{-1}$)	P_{ult} (Pa)
10^3	Rough vacuum 10^3
1	Medium vacuum 1
10^{-4}	High vacuum 10^{-4}
10^{-9}	Ultrahigh vacuum 10^{-9}

Thus, when the best possible pump is used, the main difference between a vacuum vessel which achieves only rough vacuum and one that achieves ultrahigh vacuum, is that in the latter case the total rate of influx of gas is twelve orders of magnitude less than in the former case. The other essential difference lies in the choice of the best possible pump.

The ideal surface to which all gases stick cannot be achieved. The nearest approach in practice is a cryogenic pump consisting of a surface cooled with liquid helium. This behaves in the required manner for easily condensable gases, but not for He, H_2 and Ne. There are also other limitations to such a pump, when considering gas load, etc. Indeed, all practical pumps have their limitations. For example, the range of pressure over which they function is

normally limited. It is therefore essential to select the correct type of pump or combination of pumps to achieve the desired ultimate pressure and this aspect of the problem is discussed in detail in Chapter 3. The remaining sections of this chapter present a discussion on a physical basis of the factors which determine the influx of gas into the vacuum vessel.

1.5 Sources of gas within the vacuum system

1.5.1 Leaks through the vacuum vessel

If any hole exists in the vacuum vessel, gas can pass from the atmosphere into the vacuum. It is therefore necessary to reduce the number and size of these leaks so that the total rate of gas influx is less than or equal to the product of available pump speed and the pressure. For the system already discussed in Section 1.4.3 the permissible values of Q_i are given in *Table 1.2.*

Most solid materials, when free from imperfections, are adequately leak tight. Usually leaks occur at joins in the vacuum envelope made between two parts of the same kind of material or between dissimilar materials.

For joining metals, the usual techniques of soldering, brazing and welding are all capable of providing leak-free joins. For glasses and some metal-to-glass joins, fusing of the glass is satisfactory. For materials which do not lend themselves to joining by such methods, it is often possible to make leak-tight joins using elastomer or soft metal gasket seals, or even various adhesives with suitably low vapour pressures.

The use of a helium leak detector for testing vacuum vessels provides a very sensitive test for leaks. The methods for leak detection will be discussed in Chapter 8, and therefore at this point it is only necessary to indicate that the minimum detectable influx of helium using a leak detector is of the order of 10^{-12} Pa $m^3 s^{-1}$. It can be seen from *Table 1.2* for the system previously discussed that a leak rate of 10^{-12} Pa $m^3 s^{-1}$ is already three orders less than the permissible level required to reach ultrahigh vacuum. In general, one can expect that the helium leak test is sufficiently sensitive to detect leaks which could otherwise prevent most practical systems from achieving ultrahigh vacuum.

It should be realized, however, that the stresses and strains on joins, brought about by pressure differences, vibrations and thermal cycling, can lead to failure and it is therefore most important to use methods of making joins which will remain leak tight under these conditions. The technology of making permanent seals will be discussed in more detail in Chapter 2, whilst demountable seals are dealt with in Chapter 6.

1.5.2 Virtual leaks

The nature of a virtual leak can be described most easily by a practical example. Suppose that, when assembling a vacuum system under atmospheric conditions, a screw is screwed into a blind hole, thus trapping a volume of gas at atmospheric pressure in the hole. When the vacuum vessel is evacuated, this trapped gas can leak into the vacuum system along the narrow helical crevice formed by the screw thread. The flow of gas will be controlled by the

conductance of the crevice and the pressure difference. Ultimately, all the gas will pass through and there will be no further flow of gas.

Suppose the volume of trapped gas is 10^{-6} m^3 and the conductance of the crevice is 10^{-12} m^{3}s^{-1}. The time taken to remove the gas through the hole is given approximately by

$$t = \frac{V}{C} = 10^6 \text{ s} \qquad \text{or 10 days} \tag{1.38}$$

The problem of the virtual leak is, therefore, not that it limits the ultimate pressure of the system, but rather that it may make the time required to reach the ultimate pressure very long. The actual time involved obviously depends on the magnitudes of V and C encountered. Virtual leaks can usually be avoided by proper design of components.

1.5.3 Vaporization

If the kinetic energy of any of the atoms or molecules bound together to form a solid or liquid is sufficient to overcome the binding forces, then the particle can escape into the gas phase. This process, called vaporization, may occur at any time because it is always possible, due to the random distribution of the energies of the particles, that some particles may have sufficient energy to escape.

If the bulk medium, i.e. the solid or liquid, is in a vessel which confines the surrounding gas, then all the particles which have vaporized will be trapped within the vessel and the density, and hence the pressure, of the trapped gas will rise. It has already been shown that the molecules of a gas strike a surface at a rate v per unit area given by Equation (1.9). Thus, it is to be expected that the pressure of the gas in the vessel will rise to an equilibrium value P_v at which point the rate of vaporization per unit area, W, is equal to the rate of return of the molecules. If all the molecules striking the surface actually return to the bulk then from Equation (1.9) at equilibrium

$$W = \frac{P_v}{\sqrt{2\pi mkT}} \tag{1.39}$$

The assumption that all molecules striking the surface return to the bulk has been discussed in some detail by Langmuir and also by Verhoek and Marshall (see reference 3). They conclude in the case of vapours incident on bulk surfaces of the same material, that the available experimental evidence supports the assumption.

The equilibrium vapour pressure P_v can be derived from the thermodynamic relation due to Clausius–Clapeyron, i.e.

$$\frac{dP_v}{dT} = \frac{L_B}{T(V'_G - V'_B)} \tag{1.40}$$

where V' is the specific volume of a mole of the substance and L_B the latent heat of vaporization per mole. The subscripts G and B refer to the gas and bulk phases respectively. For the gas phase under vacuum $V'_G \gg V'_B$, V'_G is given from Equation (1.1) as

$$V'_G = \frac{N_0 kT}{P_v} \tag{1.41}$$

where N_0 is Avogadro's number, i.e. the number of molecules in a mole of gas. If it is assumed that L_B is independent of the vapour pressure, Equation (1.40) can be integrated to give

$$\log P_v = a - \frac{b}{T} \tag{1.42}$$

The equilibrium vapour pressure of many substances under ultrahigh vacuum follows this law where a and b are constants. The curves of vapour pressure for different materials as a function of temperature are given in Appendix 2.

Since the rate of vaporization, W, should depend only on the number of particles having sufficient energy to escape, the value of W should be independent of the pressure above the surface. Thus the rate of evaporation into any vacuum should be constant and depend only on the temperature and nature of the substance, according to the relation

$$W = \frac{\exp\left(a - \frac{b}{T}\right)}{\sqrt{2\pi mkT}} \tag{1.43}$$

The nett flow from unit area of the evaporating substance is thus

$$W - \frac{P^1}{\sqrt{2\pi mkT}}$$

where P^1 is the pressure of the vapour above the surface. When the product of this flow and the surface area of the evaporating substance provides a gas load equal to the rate of pumping, the system will be limited by the process of evaporation.

1.5.4 Surface outgassing

In order to understand the reason for outgassing from surfaces it will be necessary to discuss the interaction of gases with solid surfaces.

A gas molecule can approach a solid surface and collide with it. It may rebound elastically but more often it will stick for a time and then leave in a direction unrelated to that from which it came. The forces which cause the molecule to stick can arise in a number of ways. The molecule may be held by the van der Waals forces betwen the gas molecule and the particles of the surface. These are the same forces which are frequently discussed as the cause of the departure of gases from the ideal gas laws. The molecules may in fact dissociate on striking the surface and the component atoms can react with the surface particles to form chemical bonds. This process could involve the migration of the molecules or atoms across the surface to sites which are more favourable for binding the gas. Whichever process occurs, the gas can be regarded as held to the surface by a bond of energy E, where E is defined as the energy liberated when one mole of the gas molecules is brought from the gas phase to the bonded state. It is often termed the activation energy of

desorption. When van der Waals forces are the only forces involved, E is always positive. In order for the gas to be released from the surface it must acquire an energy sufficient to overcome the binding energy E. Due to the thermal motion of the particles in the surface, there is always the possibility of a surface particle imparting this energy to a bonded molecule by a collision process.

According to statistics, the chance of a bonded gas molecule having energy E at any one time is proportional to $\exp[-(E/RT)]$ where R is the gas constant per mole $= N_0 k$. Thus if there are N_A molecules adsorbed, i.e. bonded on the surface, then the rate of desorption is given by

$$\frac{-\mathrm{d}N_A}{\mathrm{d}t} = BN_A \exp\left(-\frac{E}{RT}\right) \tag{1.44}$$

where B is a constant. If we assume that of all the molecules striking the surface, only a fraction, f, stick, then the rate of adsorption is given from Equation (1.9) as

$$v_A = \frac{fP}{\sqrt{2\pi mkT}} \tag{1.45}$$

f is often referred to as the sticking coefficient.

At equilibrium the rate of adsorption must equal the rate of desorption. Thus equating (1.44) and (1.45)

$$N_A = \frac{f}{B} \frac{P}{\sqrt{2\pi mkT}} \exp\left(\frac{E}{RT}\right) \tag{1.46}$$

Equation (1.46) is an expression for the amount of gas adsorbed per unit area and states that the amount adsorbed increases with the gas pressure and decreases with increasing temperature when E is positive. A plot of N_A against P at constant temperature is called an adsorption isotherm.

The assumption made in deriving Equation (1.46) may be incorrect in detail. For example, if adsorption takes place at favourable sites, then the number of sites available will be a function of the amount of gas already adsorbed. It might also happen that the presence of adsorbed molecules influences the binding energy of molecules adsorbed at a latter stage. By considering details of this kind, other workers have developed different theoretical expressions for the adsorption isotherm. The work of Polanyi[8], Langmuir[9] and Brunauer *et al.*[10] is of great importance in this respect. Whatever the form of the isotherm, it is always true that under equilibrium conditions some gas will be adsorbed on the surface. If the equilibrium conditions are changed, for example by reducing the pressure of the gas phase, then some gas must be released from the surface so that a new equilibrium condition appropriate to the new gas pressure can be achieved. Since the rate of evolution of the excess adsorbed gas will determine the flow of gas which must be pumped away, it is obviously important to investigate the rate of desorption and the time taken to reach equilibrium.

Suppose a time τ is assigned to the average time a molecule remains on the surface of a solid. Under equilibrium conditions the number of molecules adsorbed per unit area is given by

$$N_A = \nu_A \tau \tag{1.47}$$

or substituting ν_A from Equation (1.45)

$$N_A = \frac{fP}{\sqrt{2\pi mkT}} \tau \tag{1.48}$$

Comparing Equations (1.48) and (1.46)

$$\tau = \frac{1}{B} \exp\left(\frac{E}{RT}\right) \tag{1.49}$$

Equation (1.49) is more usually expressed in the form given by Frenkel[11] as

$$\tau = \tau_0 \exp\left(\frac{E}{RT}\right) \tag{1.50}$$

Since all the adsorbed gas will have undergone on average a complete change in time τ, the surface should have had the opportunity to achieve the new equilibrium condition in that time, provided, of course, that the pump speed of the system does not set a limit. Thus, the time τ can be used as an approximation for calculating the time to achieve equilibrium.

The binding energy E varies according to the type of bond and can be determined experimentally by direct calorimetry measurements. For physical adsorption the value is equal to the heat of adsorption, and increases as the boiling point increases. Thus helium, with a boiling point of 4.2 K, has a heat of adsorption of about 590 J mol^{-1}, hydrogen with a boiling point of about 20.4 K has a value of about 6.3 kJ mol^{-1}, while argon, nitrogen, oxygen and carbon monoxide with boiling points between 77.3 K to 90.1 K have heats of adsorption from 12–17 kJ mol^{-1}. For chemical adsorption, E can be much higher, for example, for oxygen on titanium, E is 1000 kJ mol^{-1}.

The value to be assigned to τ_0 represents a more difficult problem. Frenkel[11] suggested that τ_0 would be closely associated with the period of oscillation of the surface atoms in a direction perpendicular to the surface. In this case the value of τ_0 would be between 10^{-12} and 10^{-14} s. The value of τ as a function of E, is given in *Table 1.3* assuming a value of $\tau_0 = 10^{-13}$ s.

For physical adsorption of the gases mentioned, the calculated value of τ is less than 10^{-10} s at room temperature, and one could therefore infer that adsorption and desorption at a surface takes place almost instantaneously. Indeed, this is borne out in practice for a number of adsorption systems but, nevertheless, there appears to be increasing evidence to suggest that the general acceptance of $\tau_0 = 10^{-13}$ s is not valid. Values of τ_0 as large as 1 s have

TABLE 1.3. Average time of sojourn, τ, for a molecule striking a surface (from Equation (1.50))

E kJ mol^{-1}	τ *in secs at* T = *25°C*	τ *in secs at* T = *500°C*	τ *in secs at* T = *1000°C*
0.4	1.2×10^{-13}	–	–
4.2	5.4×10^{-13}	1.9×10^{-13}	1.5×10^{-13}
42	2.0×10^{-6}	6.7×10^{-11}	5.2×10^{-12}
210	9.0×10^{13} centuries	14	4.0×10^{-5}
420	–	4.0×10^{5} centuries	4.2 h

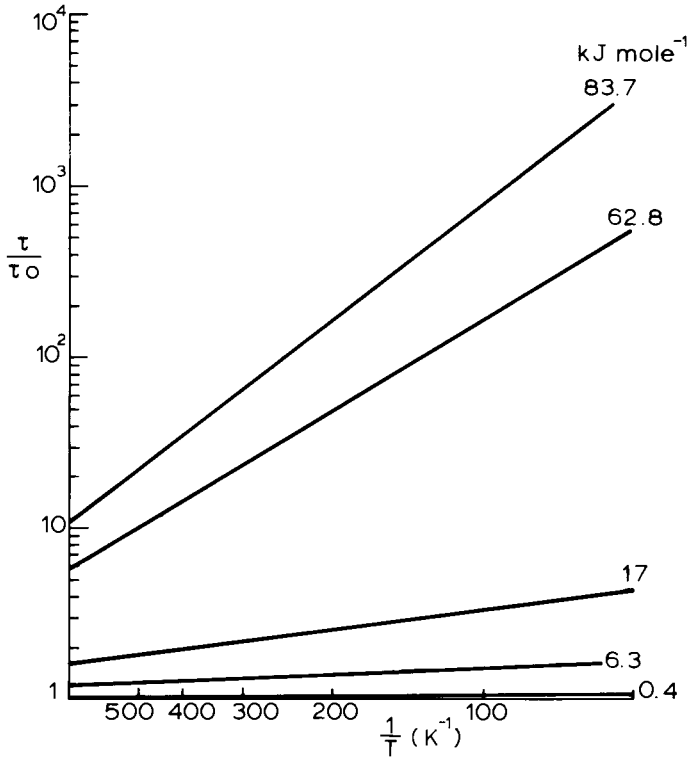
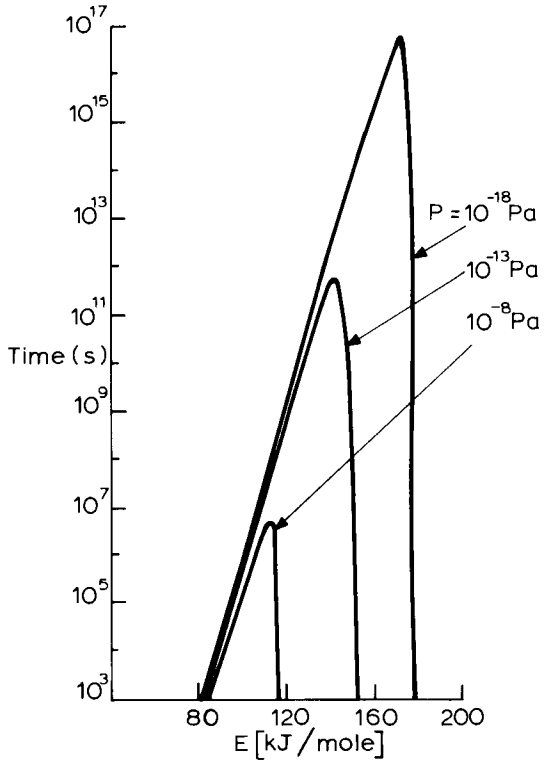


Figure 1.3 Normalized sojourn time for gas molecules adsorbed on a surface as a function of T for various activation energy values

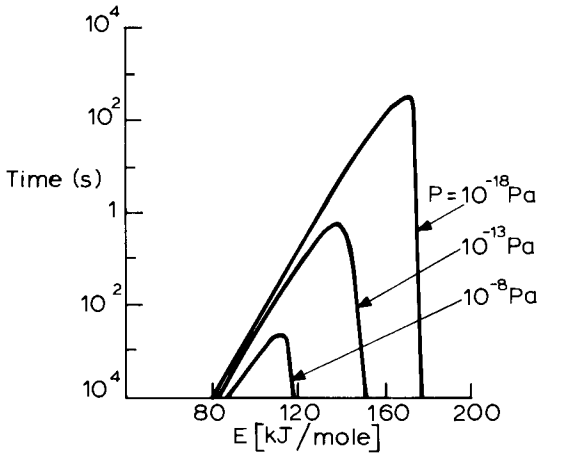
been reported. It is clear that our understanding of the processes controlling the rate of attainment of equilibrium between the gas and the adsorbed layer is far from complete and the present active research effort in this direction may provide some further insight.

In spite of the difficulty concerning the value of τ_0 it is still possible to conclude that raising the temperature of the surface will lead to a reduction in the time taken to reach equilibrium. The ratio τ/τ_0 is plotted in *Figure 1.3* as a function of temperature for different values of the adsorption energy E, showing the straight line relationship of $\log(\tau/\tau_0)$ against $1/T$ derived from Equation (1.50).

The reduction in the ratio τ/τ_0 and hence τ, brought about by raising the temperature of the surface from room temperature to 500°C, becomes significant for systems with adsorption energies exceeding 10 k cal mol^{-1}. It must be borne in mind, however, when considering vacuum systems, that it is the pump-down time to an equilibrium pressure rather than τ that is important and that for chemisorbed gases, although τ may be high, the desorption rate will be so slow as not to affect the pump down time or the ultimate pressure attained (cf. *Table 1.1*). The latter was illustrated by Redhead *et al.*[12] who plotted the time taken to reach a given pressure in an idealized system in which the gas influx was entirely due to gas desorbed for a monolayer over a surface area of 10^{-2}m^2, the pump having a speed of



(a)



(b)

Figure 1.4 Time required to reach a given pressure as a function of E according to Redhead *et al.*[12] ($A = 10^{-2}$ m^2, $S = 10^{-3}$ m^{2}s^{-1}, $V = 10^{-13}$ m^3). (a) Desorption at room temperature; (b) Desorption at 573 K

$10^{-3}m^3s^{-1}$. Their curves are reproduced in *Figure 1.4*. It is seen that the time goes through a maximum at E around 100 to 180 kJ mol^{-1} depending on the pressure considered.

Summing up, surface outgassing arises whenever the layer of adsorbed gas on a surface is not in equilibrium with the gas in the 'gas phase'. It is a transient phenomenon which ceases when equilibrium is achieved. The rate of attainment of equilibrium is, in many cases, very rapid, but circumstances may occur when the time taken is appreciable. In these cases, raising the surface temperature will secure a reduction in the time needed to achieve equilibrium.

1.5.5 Volume outgassing

In the previous section, the influence of the material below the surface was not considered. In practice, the bulk substance may contain gas which was trapped during the fabrication and treatment of the bulk substance or has diffused into it during exposure to atmospheric conditions.

For the purposes of the discussion it will be assumed that at a certain time the bulk substance is uniformly loaded with gas with a concentration c. If the concentration is disturbed so that a concentration gradient arises, then according to Fick's frst law of diffusion, a flow of gas Q per unit area will occur in the direction opposite to the concentration gradient, given by

$$Q = -D\frac{dc}{dx} \tag{1.51}$$

x being a distance from the origin of axes. D is the diffusion coefficient which is temperature dependent and usually has the form,

$$D = D_0 \exp\left(-\frac{E_d}{RT}\right) \tag{1.52}$$

where E_d is the activation energy for the diffusion process and D_0 is a constant equal to the diffusion at zero temperature.

By considering the conservation of the gas in an infinitesimal volume of the material, it follows from Equation (1.51) that the rate of change of concentration of gas at any point in the bulk substance is given by

$$\frac{\partial c}{\partial t} = D\left[\frac{\partial^2 c}{\partial x^2} + \frac{\partial^2 c}{\partial y^2} + \frac{\partial^2 c}{\partial z^2}\right] \tag{1.53}$$

for a Cartesian system of axes. This is Fick's second law of diffusion.

As an example of the application of this equation to the problem of volume outgassing, consider a slab of material with one surface exposed to the vacuum. For mathematical simplicity it will be assumed that the slab is a semi-infinite slab and initially has a uniform concentration c_0 of gas.

If the surface is exposed to the vacuum at time $t = 0$, then at the boundary $x = 0$ the concentration will be equal to the concentration in the vacuum at all times $t > 0$. Again for simplicity, a perfect vacuum with $c = 0$ at $x = 0$ will be assumed. Using these boundary conditions the solution to Equation (1.53) gives

$$c(x,t)=2c_0(\pi)^{-1/2}\int_0^{x/2(Dt)^{1/2}}\exp(-y^2)\,dy$$

$$=c_0\,\mathrm{erf}\frac{x}{2(Dt)^{1/2}} \tag{1.54}$$

The outgassing rate per unit area from the surface at $x=0$ at any time t, is then obtained from Fick's first law as

$$Q=-D\left(\frac{dc}{dx}\right)_{x=0} \tag{1.55}$$

i.e.

$$Q=-c_0D^{1/2}(\pi t)^{-1/2} \tag{1.56}$$

The general outgassing characteristics can therefore be summarized as an outgassing rate which diminishes with time according to a $t^{-1/2}$ curve and a temperature dependence of the outgassing rate through the $D^{1/2}$ term which has the form

$$D^{1/2}=D_0^{1/2}\exp\left(\frac{-E_d}{2RT}\right) \tag{1.57}$$

Raising the temperature will increase the outgassing rate and reduce the time taken to reach equilibrium when outgassing will become negligible. As in the case of surface outgassing, the degree to which raising the temperature will influence the rate of attainment of the equilibrium condition depends critically on E_d the energy of activation in this case of the diffusion process.

1.5.6 Permeation

In the previous section the flow of gas from the interior of the bulk substance by diffusion was discussed. The concentration gradient driving the diffusion was considered to be present as a result of gas dissolved within the bulk. When the material forms the envelope of the vacuum vessel, a second cause for a concentration gradient is present, due to the difference between the gas pressures on the opposite sides of the envelope material.

The amount of gas dissolved in the solid material when exposed to a gas pressure, P, usually follows Henry's law for small concentrations, i.e.

$$c=SP^n \tag{1.58}$$

where S is a constant known as the solubility and n is equal to 1 for many systems but for diatomic gases which dissociate before dissolving in metals $n=\frac{1}{2}$.

Since the pressure gradient is always present, under equilibrium conditions a steady flow of gas will pass through the envelope into the vacuum given by Fick's first law as

$$Q=-D\frac{dc}{dx} \tag{1.59}$$

Integrating by separating the variables

$$Qd = -D[c_1 - c_2] \tag{1.60}$$

where d is the width of the envelope material and c_1 and c_2 are the concentrations on the surfaces exposed to the vacuum and the atmosphere, respectively. Substituting for the concentrations using Equation (1.58) gives

$$Q = \frac{DS}{d} |P_2^n - P_1^n| \tag{1.61}$$

where DS is called the permeation constant. It increases greatly with temperature due to the exponential dependence of the factor D on temperature (see Equation (1.52)).

The permeation of gases through various materials is discussed quantitatively in Chapter 2. Generally, for most combinations of gas and solid used for vacuum envelopes, it is of no great importance at room temperatures, with the exception of the passage of helium through some glasses. However, raising the temperature of the vacuum envelope, a practice which produces advantageous effects by subsequently reducing the influence of surface and volume outgassing, will, in the case of permeation, merely serve to increase the gas flow into the vacuum system.

1.6 References

1. REDHEAD, P. A., HOBSON, J. P. and KORNELSEN, E. V., *The Physical Basis of Ultrahigh Vacuum*, Chapman & Hall, London, (1968)
2. PRESENT, R. D., *Kinetic Theory of Gases*, McGraw-Hill, New York, (1958)
3. DUSHMAN, S., *Scientific Foundations of Vacuum Technique*, Wiley, New York, (1962)
4. CLAUSING, P., *Ann Physik*, **12**, 961, (1932). [English translation *J. Vac. Sci. Technol.*, **8**, 636, (1971)]
5. COLE, R. J., *10th International Symposium on Rarefied Gas Dynamics 1976, Prog. Astronaut., Aeronaut*, **51**, Pt. 1, 261, (1977)
6. DAVIS, D. H., *J. Appl. Phys.*, **31**, 1169, (1960)
7. CHUBB, J. N., *Proceedings of the 4th International Vacuum Congress*, Inst. Physics Publication, London, p. 433, (1968)
8. POLANYI, M., *Verlandl Deut. Physik. Ges.*, **16**, 1012, (1914)
9. LANGMUIR, I., *J. Am. Chem. Soc.*, **54**, 2798, (1932)
10. BRUNAUER, S., EMMETT, P. H. and TELLER, E., *J. Am. Chem. Soc.*, **60**, 309, (1938)
11. FRENKEL, J., *Z. Physik.*, **26**, 117, (1923)
12. REDHEAD, P. A., HOBSON, J. P. and KORNELSEN, E. V., *Advances Electron. Electron Phys.*, **17**, 323, (1962)

Chapter 2

Materials for ultrahigh vacuum

2.1 Criteria for ultrahigh vacuum materials

As we saw from Chapter 1 the conditions existing within a vacuum system can truly be described as a dynamic equilibrium. The ultimate pressure which can be reached depends, on the one hand, on the effective pumping speed of the pump and, on the other hand, on the influx of gas from the vacuum envelope and any components contained within the envelope. Since there are always practical limitations to the pumping speed due to the size of the pump, cost etc., reduction of the gas influx becomes the prime objective in attaining ultrahigh vacuum conditions and sets the main criterion to the choice of materials for ultrahigh vacuum use.

The materials must have a low vapour pressure and, in order to reduce desorption, must be bakable to temperatures which are usually of the order of 450°C, without losing their mechanical strength or being chemically or physically damaged. If they form part of the vacuum envelope then permeation of gas through them must be negligible. Also they must withstand atmospheric pressure and resist corrosion when exposed to air during baking.

Gas influx and mechanical strength are not the only criteria for vacuum materials. The ease of machining or fabricating them into suitable arrangements and the ability to weld, braze or otherwise seam with leak-tight joins is also essential. In most vacuum systems there is a need to make insulated electrical connections via the walls and also, in some applications, viewing windows are required. Thus, in most vacuum apparatus, the envelope will consist of both metal and insulating materials and suitable methods of sealing the metal to the insulator must be found. Further, since a fairly wide temperature range will be experienced, the thermal expansion coefficients of the various components must be carefully matched, particularly at seals where distortion due to thermal stresses could result in leaks. Lastly, the materials chosen must be readily available at reasonable cost.

Traditionally glass has been employed for small vacuum systems, whilst the majority of large systems have been constructed of mild steel or a similar metal chosen for convenience of fabrication and cost. Glass fulfils many of the requirements for ultrahigh vacuum, provided that a suitable type is chosen with a low gas permeability. However, it is mechanically rather weak and can

be used only with softer materials such as grease, wax or synthetic rubber if demountable seals and conventional valves are required. Because such materials make it impossible to bake the system to a high temperature, an all-glass construction is only suitable if such valves and joins can be avoided. In particular, it can be utilized satisfactorily in conjunction with metal valves, etc. for small systems, especially if a pressure of 10^{-6} to 10^{-8} Pa is all that is required. For viewing ports there is really no alternative.

For lower pressures and for larger systems an all-metal construction is more suitable. Mild steel, although satisfactory for unbaked high vacuum systems, corrodes when heated in air and also shows permeation to hydrogen at the elevated temperature (see Section 2.3.3). It is therefore unsuitable for ultrahigh vacuum systems even when plated. There are other possibilities but present conditions favour stainless steels for the main vacuum envelope. They fulfil most of the stringent requirements for ultrahigh vacuum and are fairly easily obtainable at a reasonable cost.

For demountable seals and valves, stainless steels can be used in conjunction with softer metal gaskets (i.e. gold or copper) which will allow baking temperatures of 450°C to be employed. Although there are now glasses and ceramics available which can be sealed directly to stainless steel, in general the use of intermediary metals, which have been specially prepared for the purpose, is preferred for sealing the steel to insulating components.

For insulators, ceramics are superior to glass in their properties, particularly mechanical strength and the ability to withstand thermal shock. However, cost and difficulty of working often offset their advantages, so that glass is still widely used.

In the following sections the physical properties and methods of preparing these materials will be discussed with particular reference to their ultrahigh vacuum use.

2.2 Glass

The term glass can be applied to practically any compound which, after fusion, cools to a solid without crystallizing. However, it is the 'oxide' glasses which are of interest and then only those containing silica (SiO_2) as the main glass-forming constituent. Other oxides are added as modifiers to give each glass type its particular physical characteristics.

There is a variety of such glasses but for vacuum systems only certain types are used. These are basically the glasses developed in the electronic tube industry for their sealing properties to selected metals. By suitably matching the expansion coefficients, vacuum tight seals can be made between the glass and metal, provided that the metal surface can be 'wetted'.

The glasses employed, are classified traditionally into two categories, the 'hard' or borosilicate glasses, in which the main additive to the silica is boric oxide (B_2O_3), a glass-forming oxide rather than a modifier, and the 'soft' glasses in which the principal additive is either sodium oxide (Na_2O) giving soda glass, or lead oxide (PbO) giving lead glass. In *Table 2.1* the chemical composition is given of some of the glasses which are in common use in vacuum systems and which are available from Corning glass works in America. Although glasses with similar properties are available from several other manufacturers, the

TABLE 2.1. Properties of glasses commonly used for vacuum applications from Corning

Property	*Borosilicate glasses*					
	Fused silica	*Pyrex 7440*	*Tungsten sealing 7720*	*Fernico sealing 7052*	*Soda glass 0080*	*Lead glass 0120*
Chemical composition (%)						
SiO_2	100	80.8	72.2	64.3	73.2	56.2
B_2O_3		12.8	15.2	19.1		
Na_2O		4.2	3.9	5.2	16.8	3.9
K_2O			0.3		0.3	8.5
Al_2O_3		2.2	1.0	7.1	1.4	1.6
PbO			6.9			28.7
LiO				1.2		
BaO, MgO, CaO				2.7	8.2	
Viscosity-temperature characteristics (°C)						
Strain Pt.	990	515	485	435	470	395
Annealing Pt.	1050	565	525	480	510	435
Softening Pt.	1580	820	755	710	695	630
Working Pt.	–	1245	1140	1115	1005	980
Expansion Coefficient $\times 10^{-7}°C^{-1}$	5.5	33	36	46	92	89
Resistance to thermal shock $\frac{1}{4}''$ plate (°C)	1000	150	130	100	50	50
SP gravity	2.20	2.23	2.35	2.28	2.47	3.05

composition may vary slightly from the different companies and from country to country.

Apart from their chemical composition, the distinction between hard and soft glasses can be found in terms of their viscosity–temperature characteristics. Glass has no specific melting point, instead its viscosity decreases monotonically with increasing temperature until it becomes fluid. The soft glasses 'soften' and can be worked at lower temperatures than hard glasses. As a result, systems constructed with soft glass envelopes cannot be baked above 350°C without risk of deformation under the atmospheric pressure. The hard glasses on the other hand are quite safe at 400°C and some, such as Pyrex, can be baked above 500°C. Mainly for this reason the hard glasses are used for ultrahigh vacuum systems. For certain applications, such as transparency to ultraviolet radiation, or high-temperature working, other glasses may be required. If they form part of the envelope, then care must be taken to see that the properties of these glasses are also compatible with ultrahigh vacuum technology.

2.2.1 Physical properties

The physical properties of glass of importance to ultrahigh vacuum application are those affected by temperature, since this plays a vital role in the outgassing of the vacuum system. The two properties concerned are the viscosity, which is a measure of the mechanical rigidity of the glass, and the expansion coefficient, which determines the stresses and strains that are set up when uneven temperature distribution or contact with other materials occurs.

As already mentioned, glass has no definite melting or freezing point but loses its solid-like character as it is heated, by virtue of the continuous decrease in the value of the viscosity, η. The viscosity–temperature curve depends on the composition of the glass and in *Figure 2.1* some typical curves of $\log \eta$ against temperature are given for glasses of similar composition to those described in *Table 2.1*.

Four values of viscosity on the viscosity–temperature curve have been defined by the American Society for Testing Materials (ASTM) to represent the texture of the glass as it changes from solid to liquid, and these are generally accepted internationally. The temperatures corresponding to these four values are designated the strain point, the annealing point, the softening point and the working point. The strain point represents the temperature at which internal stresses are relieved in a few hours, defined by a viscosity value of $10^{13.5}$ Pa s^{-1} ($10^{14.5}$ poise). The annealing point is the temperature at which internal stresses are relieved in a matter of minutes, defined by a viscosity value of 10^{12} Pa s^{-1}. The softening point is defined in terms of the elongation of a standard fibre under its own weight and corresponds to the temperature at which the viscosity is $10^{6.6}$ Pa s^{-1} for glasses with a density around 2.5 g cm^{-3}. The working point is the temperature at which the glass is soft enough to be worked by normal fabricating techniques and is defined by a viscosity value of 10^{3} Pa s^{-1}. These temperature points have been inserted in *Table 2.1* for the various glasses. The temperature at which the vacuum envelope would deform under atmospheric pressure, depends on the shape and thickness of the glass and the length of time

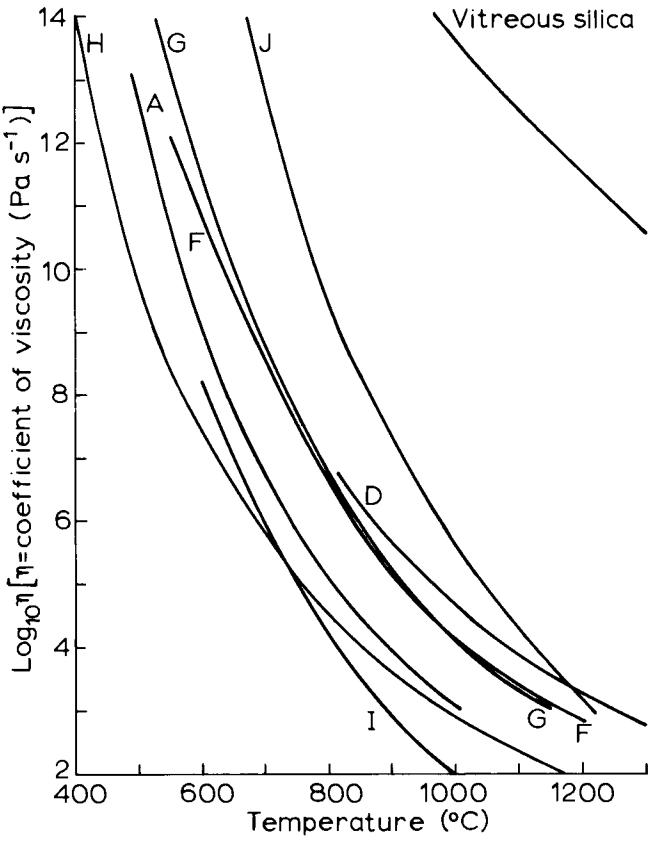


Figure 2.1 Viscosity–temperature curves for various glass types according to Douglas[1]. A, soda glass; D, hard borosilicate glass (similar to Pyrex); F, tungsten sealing borosilicate; G, an alternative tungsten sealing borosilicate; H, lead glass; J, alumino–silicate glass 54% SiO 21% Al_2O_3, 8% B_2O_3; I, soda vapour resistant glass 23% SiO 24% Al_2O_3 37% B_2O_3

it is held at the temperature but as a guide, the strain point can be taken as the upper temperature limit for safe bakeout.

The change in viscosity with temperature, however, is of less significance than the effect of thermal expansion. Glass when heated tends to expand and although the expansion is relatively small compared with other materials, the effect can cause stresses and strains within the glass which, because of its brittle nature, can result in fracture. It is of particular concern in connection with rigid seals between the glass and other materials, such as metal and ceramics, and also where dissimilar glasses are joined. Some typical expansion curves expressed as fractional elongation against temperature are shown in *Figure 2.2*. In general, the expansion is greater for the softer glasses and in the borosilicate glasses decreases as the amount of B_2O_3 decreases. Below 300°C the curves are essentially linear and a constant expansion coefficient can be ascribed to the glass; values of this expansion coefficient for the various glasses are given in *Table 2.1*. At higher temperatures the expansion coefficient

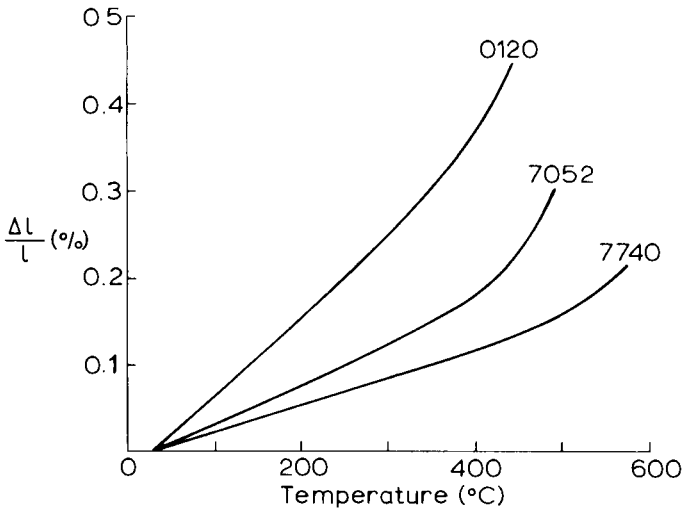


Figure 2.2 Expansion–temperature curves for three Corning glasses

increases and it becomes appreciably higher as the annealing temperature is approached. The values of the expansion coefficient are only reproducible and reversible for well-annealed glasses. Poorly annealed and strained glass will give somewhat higher values of the coefficient and exhibit irreversible irregularities in the expansion v. temperature curve.

Stresses are not only set up at seals as a result of expansion, thermal gradients across the glass can also be a danger. If one side of a glass plate is hotter than the other, then the heated side will experience a compressional stress whilst the cooler side will be under tension. The latter causes the failure in glass. The tensile stress set up in this case will depend on the temperature gradient and on the glass properties, particularly the coefficient of expansion. In general, the lower the expansion coefficient the higher will be the temperature gradient which the glass can withstand. Thus, for a constrained plate, the temperature difference causing a tensile stress of 7×10^6 Nm^{-2} is about 50°C for Pyrex, but only about 15°C for soda glass. Although large steady temperature gradients are not likely to occur in practice, large transient gradients may well occur, for example when a glass condensation trap is first immersed in liquid nitrogen. The strength of glass is greater under momentary stress than under prolonged stress, so that the resistance to thermal shock cannot be assessed from the static characteristics. It depends not only on the expansion coefficient but also on the shape of the sample, its thickness and whether the stress is incurred by sudden heating or cooling (the latter is the more damaging).

An empirical testing schedule used by Corning, where a plate of given dimensions is heated and then plunged into cold water, gives an indication of the resistance to thermal shock. The highest temperature to which the plate can be heated, without damage on cooling, is taken as the criterion, and the values for the Corning glasses have been inserted in *Table 2.1*. In general, soft glasses are unsatisfactory as cooling traps or other parts of the vacuum

envelope subject to thermal shock. Its low price and ease of working, however, has seen its widespread use for electronic tube envelopes.

2.2.2 Permeation of gases through glass

From the early experiments in vacuum, it was known that gas could permeate through a thin glass wall and there are several references[2] to the measurement of the permeation of gas through silica and glass in the 1920s and 1930s. However, it was considered that, for practical purposes, the rate at which gas 'leaked' into a vacuum system or an electronic tube from the atmosphere was so small at room temperature that the effect could be ignored.

With the attainment and measurement of ultrahigh vacuum in glass systems, permeation of gas through the walls from the surrounding atmosphere was recognized as a contributory source of gas influx, limiting the ultimate pressure. For example, in 1954 Alpert and Buritz[3] reported that, in their Pyrex glass system, permeation of atmospheric helium (the equilibrium pressure of helium in air is $\sim 5.3 \times 10^{-1}$ Pa) through the walls was the predominant source of residual gas. They observed that, in a sealed-off volume of 400 l, the pressure rose from 2×10^{-7} to 2×10^{-6} Pa in about 10 h.

The microstructure of glass has the general form of SiO_4 tetrahedra sharing oxygen atoms, as in crystalline quartz, but in an irregular arrangement making a more open structure in which gas atoms can be sited, see *Figure 2.3*. The addition of modifiers Na^+ and K^+, etc. occupies some of the sites within the pockets surrounded by the silica or silica–borate structure. Thus, it might be

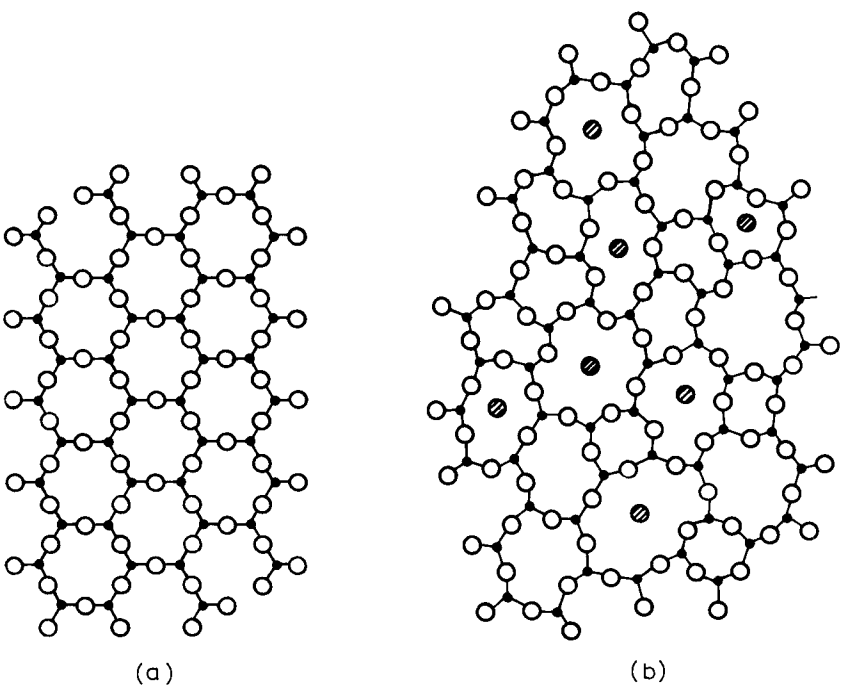


Figure 2.3 Two-dimensional schematic representation of the structure of (a) quartz crystal and (b) glassy form of silica with modifier atoms shown as hatched circles

expected that the permeation of gas through glass will depend on the 'porosity' of the microstructure and be reduced by the presence of modifiers. It can also be expected that the size of the gas molecules will be important.

The quantity, Q, of gas permeating through a solid wall or membrane of thickness d and area A is given by Equation (1.61) namely

$$Q = \frac{Ak}{d}(P_2^n - P_1^n) \tag{2.1}$$

where P_1 and P_2 are the gas pressure each side of the wall and K is the permeation constant.

For glass where n is found to be 1

$$Q = \frac{AK\,\Delta P}{d} \tag{2.2}$$

Using SI units with Q expressed as Pa $m^3 s^{-1}$ and P as Pa, both normalized to 25°C, and A and d expressed in metres, K will be given in $m^2 s^{-1}$*. Since K depends on the diffusion constant, the value of K increases exponentially with temperature according to the following equation

$$K = K_0 \exp\left(\frac{-E}{RT}\right) \tag{2.3}$$

where E is the activation energy, R the gas constant and K_0 a constant of proportionality (cf. Equation (1.52)). Thus it is convenient to present data on permeability by $\log K$ against $1/T$ plots.

Measurements of permeation confirm the general premises mentioned above. The effect of gas molecular diameter is demonstrated in *Table 2.2*, where the permeation constant K for gases passing through fused silica at 700°C, according to Norton,[4] is listed against the relevant atomic or molecular diameter. It is seen that helium with the smallest atomic diameter has the highest permeation rate, whilst argon, nitrogen and oxygen are too large to permeate appreciably. For practical purposes the silica can be considered as impervious to the latter gases. The results on hydrogen compared with neon,

TABLE 2.2. The atomic or molecular diameter in Ångström units and the permeation constant for gases through fused silica at 700°C (Norton[4])

Gas	*Atomic/molecular diameter* (Å)	*Permeation constant*, K ($m^2 s^{-1}$)
Helium	1.95	1.7×10^{-11}
Neon	2.4	3.5×10^{-13}
Hydrogen	2.5	1.7×10^{-12}
Deuterium	2.55	1.4×10^{-12}
Oxygen	3.15	Less than 10^{-18}
Argon	3.2	Less than 10^{-18}
Nitrogen	3.4	Less than 10^{-18}

* K is often expressed as cm^3 as STP per second across 1 mm thickness per cm^2 area with a pressure difference of 1 cm Hg (cm^3 (STP mm $s^{-1} cm^{-2} cm - Hg^{-1}$). To convert to SI units at 23°C multiply by 8.24×10^{-4}

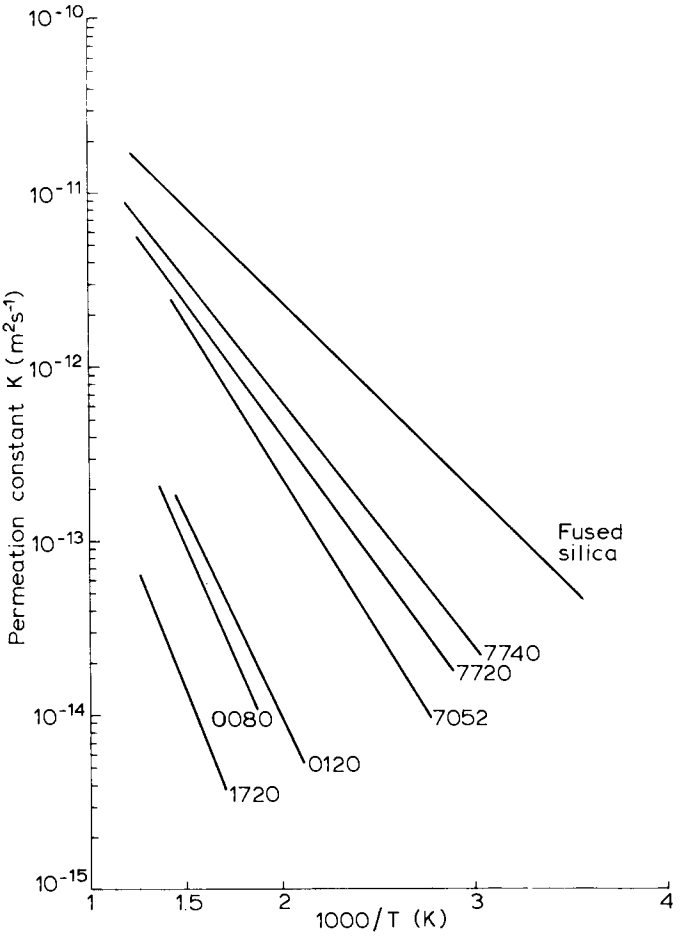


Figure 2.4 Permeation constant for helium through some Corning glasses as a function of temperature (from Altemose[5])

however, indicate that atomic diameter is not the only factor. Norton suggests that the greater permeation rate for hydrogen is related to surface and solubility effects.

The permeation rates of helium through glasses of different compositions have been investigated by several workers, and give general agreement. The values of K as a function of $1/T$ plotted in *Figure 2.4* are taken from the data of Altemose[5] converted to SI units. The figure shows the values for Corning glasses listed in *Table 2.1* and also for Corning 1720, a special aluminosilicate glass particularly suitable for high vacuum as far as permeation is concerned.

The figure shows that the permeation rates for all the glasses are lower than for silica and thus permeation of gases other than helium can be ignored. In general, the permeation rate decreases as the percentage of glass network formers SiO_2 and B_2O_3 decreases and correlation between K and the weight percent of glass former and also the glass density was demonstrated by Norton[4]. However, as Altemose[5] pointed out, such a correlation does not

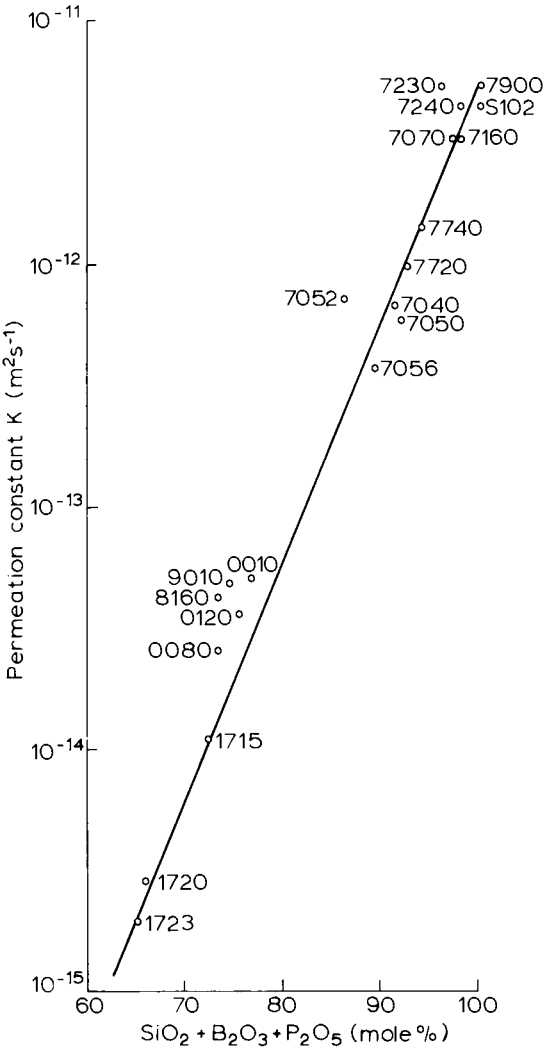


Figure 2.5 Permeation constant for helium through some Corning glasses at 300°C as a function of the mole percent network formers (from Altemose[5])

satisfy the results for lead and soda glass. He suggested that mole percent rather than weight percent should be used on the grounds that it was the packing density of the atoms rather than their mass which was the controlling factor. The plot of $\log K$ against mole percent of $SiO_2 + B_2O_3 + P_2O_5$* for a large selection of Corning glasses according to Altemose is shown in *Figure 2.5*.

The influx of helium from the atmosphere and the effect on the ultimate pressure can be calculated from the value of K. To illustrate the effect, the rise in pressure in a sealed-off system, assuming equilibrium conditions, is plotted

* P_2O_5 is a network former added to some of the glasses

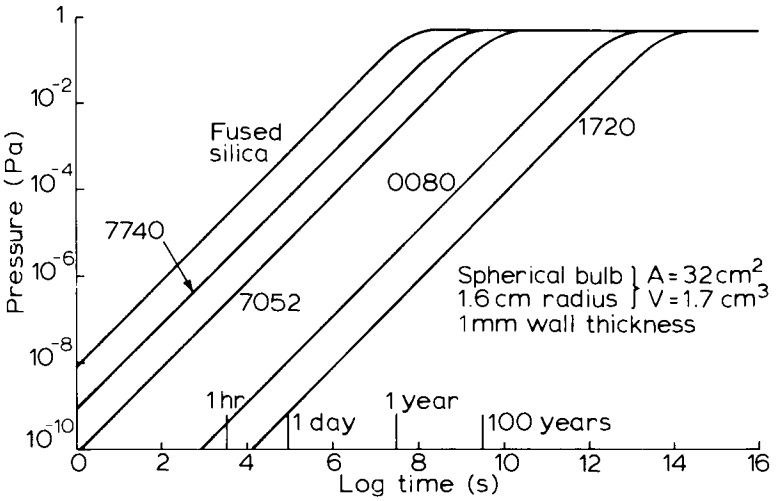


Figure 2.6 Accumulation of helium in a glass vacuum bulb as a result of permeation from the atmosphere at 25°C for various glasses

in *Figure 2.6* as $\log P$ against $\log t$ for a 1.6 cm radius bulb, with a 1 mm wall thickness, made of various glasses. Thus, at room temperature the pressure in such a bulb made of silica will rise to 10^{-5} Pa in a matter of minutes, for Pyrex it will take a few hours, for molybdenum sealing borosilicate glass, a few days and soda glass, several years. Pyrex is not, therefore, a suitable glass for ultrahigh vacuum systems, but it should be pointed out that if the temperature is raised to say 400°C, then even for soda glass in the example given, a similar pressure rise will occur in less than an hour.

2.2.3 Outgassing of glass

The permeation of gas through glass implies a certain solubility of the gas in the glass, which will diffuse into the vacuum system at a rate depending on the concentration and temperature. Gas trapped in the open network structure of the glass between the constituent atoms is said to be physically dissolved. However, in glass there is also the possibility of chemical solubility, whereby the presence of gases in the glass is due to chemical reaction. In this case larger molecules can be dissolved and also the quantity can be greater. Infrared spectroscopy and other techniques have shown that water vapour, carbon dioxide, oxygen and sulphur dioxide are dissolved in this way, the gases being dissolved during the manufacturing process whilst the glass is in a molten state. Water vapour constitutes the major component of dissolved gas, with a solubility some two orders higher than helium. Unlike helium or other physically dissolved gases, the quantity of water vapour dissolved in glass is greater the higher the percentage of alkali modifiers; consequently the solubility in soda or lead glass is considerably higher than in borosilicate glass.

In addition to the gas dissolved within the glass there will also be the gas adsorbed on the surface. Again water vapour is the main constituent, which appears to be bound fairly tightly probably as surface hydrates.

The adsorbed and dissolved gases constitute a source of gas influx in glass

envelope systems which would prohibit the attainment of ultrahigh vacuum in an unbaked system; Alpert[6] quotes a figure for the gas influx from an unbaked borosilicate glass system of 10^{-5} Pa $m^3 s^{-1} m^{-2}$. The gas influx, however, can be enormously reduced by an outgassing bake. Early investigations established that as glass is heated in vacuum a rapid evolution of gas occurs at temperatures around 200°C–300°C. Further, heating produces a slower but more persistent evolution of gas. Both gas evolutions are predominantly water vapour and it is generally accepted that the major part of the initial evolution is due to surface adsorbed layers whilst the more persistent evolution is due to gas diffusing from the interior.

A general review of the measurements made on gas desorption from glass is to be found in Dushman[2]. Of interest are the studies of Todd[7], who substantiated the general hypothesis by showing that, after an initial period, the gas evolution is inversely proportional to the square root of the baking time in accordance with a diffusion process (cf. Section 1.5.5). He further established that the constant of proportionality is exponentially dependent on temperature and also depends on the glass composition. An indication of the gas evolution from different glasses can be seen in *Table 2.3*, where the amount of gas evolved from within the glass during various bakeout conditions has been calculated from data presented by Todd[7]. The amount is expressed in Pa $m^3 m^{-2}$, but if one considers a litre volume then the figures would be equivalent to the pressure rise in Pa assuming that the gas is coming from 10 cm^2 of glass surface.

TABLE 2.3. Gas evolution from glass calculated from data by Todd[7]

Glass type (Corning)	*Gas evolved during 1 h bake in Pa $m^3 m^{-2}$*		*Gas evolved during 10 h bake in Pa $m^3 m^{-2}$*	
	at 600 K	*at 800 K*	*at 600 K*	*at 800 K*
7740	8.3×10^{-4}	6.2×10^{-2}	2.6×10^{-3}	2.0×10^{-1}
7720	1.5×10^{-3}	8.4×10^{-2}	4.7×10^{-3}	2.7×10^{-1}
0080	1.8×10^{-3}	3.3×10^{-1}	5.7×10^{-3}	1.05
0120	6.5×10^{-3}	2.8×10^{-1}	2.1×10^{-2}	9.2×10^{-1}
1720	2.0×10^{-6}	2.1×10^{-3}	6.3×10^{-6}	6.7×10^{-3}

By baking a glass system at high temperature for a period of 24 hours, the surface adsorbed layers of gas and sufficient absorbed gas are driven out to reduce the subsequent outgassing rate at room temperature to the order of 10^{-12} Pa $m^3 s^{-1} m^{-2}$.

2.3 Metals

Metals can be used in ultrahigh vacuum systems as internal components for their electrical, mechanical, or heat conducting properties, or as an integral part of the vacuum envelope. As internal components, the specific use will dictate the most suitable choice of metal. Thus for grids, where fine wires of high mechanical strength are required, molybdenum or tungsten is employed, whereas for plate electrodes requiring complex shapes, softer metals which are more easily fabricated, such as nickel and iron, are more appropriate. As far as

the vacuum requirements are concerned, the metal for internal components only has to satisfy two conditions; the vapour pressure of the metal at its working temperature must be below the desired ultimate pressure and also the metal must be capable of adequate degassing so that it does not constitute a major source of gas influx. When using the metal as an integral part of the vacuum envelope, gas from the surrounding atmosphere must not permeate through it and it should not suffer corrosion during bakeout periods.

2.3.1 Vapour pressure of metals

Honig[8] has compiled vapour pressure data from the literature, which covers most of the common elements down to the ultrahigh vacuum pressure range and some of the curves are given in Appendix 2. For convenience, values have been extracted from Honig's data for those metals either used in vacuum or likely to occur as impurities and these are given in *Table 2.4*. The data are presented as temperature values of the respective metal corresponding to the given vapour pressure.

At room temperature few metals have vapour pressures above 10^{-9} Pa and although care should be taken to avoid materials which may contain impurities such as sodium or potassium, the selection of metals for components is not seriously restricted on this score. One high vapour pressure metal, however, which is often overlooked is cadmium, and for ultrahigh vacuum systems, cadmium plated screws, etc. should not be employed.

As the temperature of the system is raised, other metals will have vapour pressures above 10^{-9} Pa and, since it will normally be necessary to outgas the system at elevated temperatures, up to 450°C, zinc, lead, tin and similar metals should be excluded. If the temperature of the component in its application is raised to higher temperatures, for example for thermionic emission, then the selection of the metal is more limited. Even a tungsten filament emitter running at 2000°C will have a vapour pressure of 10^{-7} Pa in its vicinity which, as Alpert and Buritz[3] have pointed out, could represent a pressure limiting value in an ion gauge of 10^{-10} Pa, due to the presence of the tungsten atoms. Nevertheless, for most vacuum applications there is a wide range of metals whose vapour pressures are low enough, even at elevated temperatures, not to constitute a problem down to 10^{-11} Pa pressure, many of which are readily obtainable and are economic to use.

2.3.2 Outgassing of metals

The outgassing properties of metals warrant more careful consideration, not because metals differ greatly in their outgassing rates, but because of the processing necessary to reduce the outgassing rate to an acceptable level. As with glass, gases are adsorbed on the metal surfaces and also dissolved within the metal, mainly during the manufacture and processing of the raw material. The gases may be physically or chemically sorbed, both in the case of surface adsorption and absorption throughout the bulk. When the metal is placed in a vacuum environment these gases will be evolved, at a rate depending on the total gas sorbed, the nature of the sorption process and the temperature. In

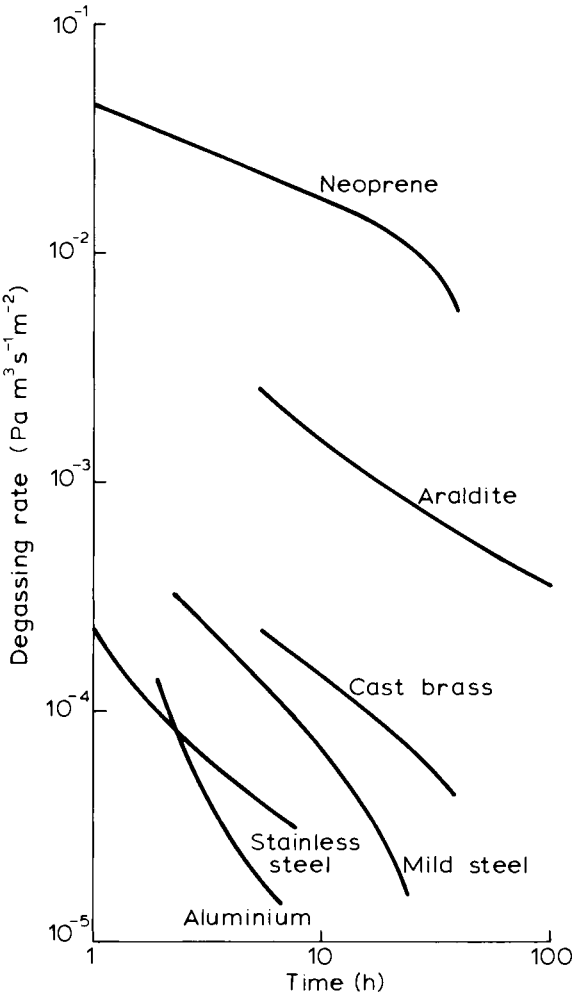


Figure 2.7 Degassing rate of untreated metals compared to Neoprene and Araldite according to Blears *et al.*[9]

particular, the activation energy of the sorption process will be important*. For untreated metals, particularly if the surface is contaminated by a thin layer of oxide, the initial outgassing rate at room temperature is of the order of 10^{-4} Pa $m^3 s^{-1} m^{-2}$, i.e. an order of magnitude greater than found for untreated glass. *Figure 2.7* gives the outgassing rates for some of the (untreated) metals used in the fabrication of vacuum systems, as measured by Blears *et al.*[9] Values for neoprene and Araldite are given for comparison. The outgassing rate decreases with time, and according to Dayton[10] can be

* If gas is chemisorbed with a high activation energy it may not be released, and indeed metals exhibiting strong bonds with the active gases, if freshly formed or cleaned in vacuum, can be employed for reducing the pressure in the system (cf. Section 3.6)

TABLE 2.4. Vapour pressures of metals expressed as temperature required to give the vapour pressures indicated

Metal	*Melting point*	*Temperature* (K) *Giving* $P = 1.33 \times 10^{-9}$ Pa	$P = 1.33 \times 10^{-8}$ Pa	$P = 1.33 \times 10^{-7}$ Pa	$P = 1.33 \times 10^{-6}$ Pa	$P = 1.33 \times 10^{-5}$ Pa
Ag Silver	1234	721	759	800	847	899
Al Aluminium	932	815	860	906	958	1015
Au Gold	1336	915	964	1020	1080	1150
Ba Barium	983	450	480	510	545	583
Be Beryllium	1556	832	878	925	980	1035
C Carbon	–	1695	1765	1845	1930	2030
Ca Calcium	1123	470	495	524	555	590
Cd Cadmium	594	293	310	328	347	368
Ce Cerium	1077	1050	1110	1175	1245	1325
Co Cobalt	1768	1020	1070	1130	1195	1265
Cr Chromium	2176	960	1010	1055	1110	1175
Cs Caesium	302	213	226	241	257	274
Cu Copper	1357	855	895	945	995	1060
Fe Iron	1809	1000	1050	1105	1165	1230
Ge Germanium	1210	940	980	1030	1085	1150
Hg Mercury	234	170	180	190	201	214
In Indium	429	641	677	716	761	812
Ir Iridium	2727	1585	1665	1755	1850	1960
K Potassium	336	247	260	276	296	315
La Lanthanum	1193	1100	1155	1220	1245	1375
Mg Magnesium	923	388	410	432	458	487
Mn Manganese	1517	660	695	734	778	827
Mo Molybdenum	2890	1610	1690	1770	1865	1975
Na Sodium	371	294	310	328	347	370
Ni Nickel	1725	1040	1090	1145	1200	1270
Pb Lead	601	516	546	580	615	656
Pd Palladium	1823	945	995	1050	1115	1185
Pt Platinum	2043	1335	1405	1480	1565	1655
Re Rhenium	3463	1900	1995	2100	2220	2350
Rh Rhodium	2239	1330	1395	1470	1550	1640

Sb	Antimony	903	447	498	526	552	582
Se	Selenium	490	286	301	317	336	356
Sn	Tin	505	805	852	900	955	1020
Sr	Strontium	1043	433	458	483	514	546
Ta	Tantalum	3270	1930	2020	2120	2230	2370
Th	Thorium	1968	1450	1525	1610	1705	1815
Ti	Titanium	1940	1140	1200	1265	1335	1410
W	Tungsten	3650	2050	2150	2270	2390	2520
Zn	Zinc	693	336	354	374	396	421
Zr	Zirconium	2128	1500	1580	1665	1755	1855

expressed over the first ten hours or so by the equation

$$q_t = q_0 / t_h^n \tag{2.4}$$

where q_t is the outgassing rate at time t_h (in hours), q_0 is about 10^{-4} Pa $m^3 s^{-1} m^{-2}$, and n may vary from 0.7 to 2 but is frequently in the neighbourhood of 1.

As with glass, this large outgassing rate, which is mainly due to water vapour, is ascribed to gas physically adsorbed on the metal surface. However, from the total quantity of gas evolved, several monolayers are involved; some experiments suggest more than a hundred monolayers. In the case of metals having a more or less porous layer of oxide on the surface, the gas is likely to be sorbed throughout the oxide film. It can then be expected that the degassing rate will be controlled by diffusion of the gas through the pores or grain boundaries of the coating. This can explain why the time of removal is prolonged compared with glass and with the theory. Also it may be expected that the surface outgassing will be affected by the surface condition. Reiter and Camposilvan[11] found significant differences in outgassing rates between mechanically polished and sand-blasted Inconel 600. The sand blasting increased the surface area but on the other hand broke up the oxide surface layer. The result was an increase in the outgassing rate for those gases adsorbed or formed on the surface (see p. 39) such as CO_2 and CO and a decrease in the degassing rate for those gases such as H_2 which had to diffuse through the oxide barrier.

After the first 10–100 hours when the bulk of the surface adsorbed gas has been evolved, the degassing rate decays exponentially to a much lower value, when the gas diffusing from the bulk material becomes the controlling factor, i.e. the rate is then inversely proportional to the square root of the time. The gases within the bulk metal are commonly H_2, N_2, O_2, CO and CO_2, and not water vapour. They are mostly taken up during the molten stage and arise from the furnace gases during the melting process and/or from ambient gases during casting. Some of the gases listed above react with the metal to form compounds, whilst others are physically held in solution. The mechanism will depend on the metal–gas combination. When the melt solidifies the gases are only partially released leaving a considerable gas volume still entrapped in the metal. Indeed, typically the gas content of metals is some 10–100% of their volume at STP, which is of the same order as the water vapour content of glass.

The gas content of metals is not always attributable to the melting process, since gases can diffuse into most metals in the solid state. The process is relatively slow at room temperature, and metals which have been degassed can generally be stored for periods of days without an appreciable increase in gas content. The diffusion however increases exponentially with temperature. The rare gases are an exception and do not dissolve in any metal under purely thermal conditions, even when the metal is molten. They can only be absorbed if the gas atoms bombard the metal with a high energy, either as ions or energetic neutrals.

A considerable amount of information exists on the solubility of gases in metals, mainly in the solid phase; a summary can be found in Dushman[2]. Some general rules which can be applied when considering the solubility of gases in metals are given below under the headings of the gases involved.

Hydrogen

For hydrogen solubility, the metals can be divided into two main groups; those which form solid solutions and those forming hydrides or pseudo-hydrides. In the first group the solubility is proportional to the square root of the partial pressure of hydrogen and increases with temperature. Metals in this group listed roughly in order of increasing solubility are Al, Cu, Pt, Ag, Mo, W, Cr, Co, Fe and Ni. In the second group, the metals forming true hydrides are the alkali and alkali–earth metals, i.e. Na and Ca, etc. and also elements of group IVb, Vb and VIIb such as B, C, S, Si and As. The so-called pseudo-hydrides have the form MH_m where m is not an integer, and include the metals Mn, Ta, V, Nb, Ce, La, Zr and Ti. Their hydrogen solubility can be orders of magnitude greater than that for metals in the first group and they are of interest for their gettering properties. The solubility of the second group decreases with increasing temperature.

Oxygen

Oxygen is soluble to some extent in most metals but, except in the case of the noble metals, an oxide phase also appears when the limit of solid solubility is exceeded. It is difficult therefore to distinguish between oxygen from solution and from the oxide phase. During the molten state, many metals will take up large quantities of oxygen, which are precipitated as oxide when the metal solidifies. As a result, the quantity of oxygen in the metal, capable of being released into a vacuum system, can be large compared with the amount which can be dissolved in the solid metal.

Nitrogen

Nitrogen dissolves only in those metals which form nitrides at higher temperatures, for example Zr, Ta, Mn, Mo and Fe. It has been shown to be insoluble, within experimental limits, in Co, Cu, Ag and Au. For the metals like Mo and Fe, the solubility is small, of the order of 1% or less by weight, and there is no tendency to form nitrides in the solid phase. On the other hand, Zr when heated, dissolves large quantities of nitrogen to form the nitride.

Carbon monoxide and dioxide

A similar behaviour to nitrogen is encountered with CO. It is dissolved only by metals such as Ni and Fe, which are capable of forming a carbonyl. CO is not soluble in copper. The observed evolution of carbon monoxide and dioxide on heating some metals in vacuum may be accounted for by diffusion of carbon and its reaction with the oxides in the metal and not direct solution of CO or CO_2.

Metals can be degassed in a similar way to glass by baking in the final vacuum system. However, since the rate of diffusion is exponentially dependent on temperature, for the dissolved gas, it is advantageous to go to the highest temperatures that metals are capable of withstanding, rather than the most convenient for the vacuum system as a whole. This outgassing can be carried out before assembly in the vacuum system, since, as has already been stated, re-absorption of gas at room temperature is a relatively slow process.

Thus, the outgassing of metal parts can be carried out in stages, used alone or more usefully in combination, as follows:

(1) Melting the raw material in vacuum to provide gas free ingots.
(2) Heating the preformed parts in vacuum before assembly.
(3) Heating the assembled parts in the final vacuum system.

The vacuum melting of metals can now be carried out on an industrial scale and most metals are available in a vacuum-melted form. Details of the techniques and the apparatus used for the degassing of the crude metal during melting are to be found in the comprehensive book by Espe[12]. However, gas-free vacuum melted metals are relatively expensive and are therefore normally limited to applications where freedom from oxygen or deoxidant impurities are essential, e.g. cathode nickels and glass-sealing alloys. For the majority of vacuum applications there is little to be gained from using gas-free ingots, since, in fabricating the components, it is difficult to avoid contamination by oil, etc. or to avoid oxide formation, especially if hot-forging, hot pressing or welding are involved.

The degassing of the fabricated components prior to their assembly, on the other hand, is almost essential. The fabricated components are first chemically cleaned to remove surface oxide or other contaminant layers and degreased to remove oil, which may be deposited on the surface during handling. The component is then heated in a vacuum of 10^{-2} to 10^{-3} Pa to a temperature of the order of 1000°C. Alternatively the component can be heated in a flushing gas which is either not readily absorbed or easily removed at a later date. For example, hydrogen stoving is used extensively for electron tube parts. The hydrogen reduces any oxide present and, because of its high diffusion rate, is driven off fairly readily during the subsequent tube processing. It is also a more economic stoving process as the absence of vacuum seals makes it possible to feed the components continuously through the heated furnace on a moving belt system. However, the removal of the residual hydrogen, which is not detrimental in an electron tube, can be a problem in ultrahigh vacuum systems, so that hydrogen stoving is not recommended, even when followed by vacuum stoving.

The actual temperature of the stoving furnace and the stoving time are dictated by several factors. Clearly the temperature must not be so high that appreciable evaporation of the metal takes place, or that the melting point is approached. In practice, a lower limit is set by distortion of the component (creep limit) or in some cases by the design of the stove. In general the temperature will be about the annealing temperature and as a result of the stress equalization the component may warp and have to be trued up later. In the case of stressed parts such as supports and springs, the temperature must be kept below the annealing temperature. For most metals, temperatures from 900–1000°C are adequate and these can be attained with resistance wound furnaces employing nichrome or Kanthal windings. Higher temperatures are desirable for tungsten and molybdenum, requiring more elaborate furnaces; Norton and Marshall[13] found that nitrogen, which formed more than 50% of the sorbed gas in molybdenum was not released until temperatures above 1200°C were reached. Titanium and copper must be stoved at lower temperatures, 500–700°C, and only oxygen-free high conductivity (OFHC) grade copper is suitable for high-vacuum applications. Aluminium, because of

its low melting point, cannot be stoved much above the normal bakeout temperature of the completed system and, therefore, there is little to be gained by a pre-outgassing process. Hydrogen is not necessary to reduce the metal oxides since most of the metal oxides will decompose at the stoving temperature, giving off oxygen. The exceptions are the few oxides such as Al_2O_3, MgO and ThO_2 having low dissociation pressures.

The stoving time should be as long as practicable but, since most of the gas is removed in the first few hours, a stoving time of eight hours at the equilibrium temperature is usually sufficient. Taking into account the heating and cooling time, this allows stoving to be carried out overnight.

As a result of vacuum stoving, the outgassing rate in the final vacuum system may be reduced by several orders of magnitude. Flecken and Nöller[14] observed an 88–97% reduction in the amount of gas desorbed after stoving stainless steel at 800°C for 1 h followed by exposure to the atmosphere for 1 h. Care must be taken in storing and handling the components between stoving and assembly, since as Varadi[15] has shown, touching the component with fingers can markedly increase the degassing rate and surprisingly some of the increase could be attributed to gas re-absorbed into the bulk metal.

Even with careful handling, surface contamination can occur and it can be expected that gas adsorption will take place on the surface of the stoved parts. Thus a further degassing process *in situ* is necessary to achieve ultrahigh vacuum pressures. Also for small components, such as filaments and grids, higher temperatures can be reached in the vacuum system than can be conveniently achieved in a vacuum furnace. Heating the metal in the vacuum system for degassing purposes can be carried out in one of four ways: (1) baking in an external furnace; (2) direct passage of current through the component; (3) eddy current heating by HF induction; (4) electron or ion bombardment.

Method (1) is generally limited to temperatures below 500°C, whilst methods (2), (3) and (4) allow much higher temperature outgassing to be utilized. Eddy current heating is restricted to components mounted in an envelope of glass or other insulating material but nevertheless is a useful method of outgassing sheet metal electrodes to high temperatures in experimental electronic devices; for details see Espe[12]. Heating by direct passage of current can be applied to filaments and components of fairly high resistance, for example helical wound grids. The component should be heated to a temperature in excess of that reached during subsequent operation. Electron bombardment of a component is probably the most versatile method of heating metal parts to temperatures above 500°C. Provided a suitable electron source is available, it imposes no restriction on the target material or its shape and can be conveniently controlled by circuitry. Ion bombardment can also be used. It not only heats the target, but also removes the surface by sputtering. In applications where surface cleaning is required it can be used with advantage, although usually a gas pressure above 10^{-2} Pa is required to obtain the necessary ions.

Ion sputtering using an inert-gas glow discharge can be used for the treatment of metal envelopes, particularly those used for particle accelerators. Although Govier and McCracken[16] reported rather variable results from such treatment, Jones *et al.*[17], who carried out the process whilst the envelope was being baked at 300°C, reported that the quantity of gas released subsequently

by electron bombardment was 15 times lower than obtained by baking alone. It is now fairly commonly used for large accelerator envelopes and a review on glow-discharge cleaning methods has been given by Holland[18].

Methods (2), (3) and (4) are normally restricted to relatively small components. For large components, especially the metal vacuum envelope, baking with an external furnace provides the only feasible method of thermally outgassing the assembled system. Since much of the surface-adsorbed gas is water vapour, baking at 200–250°C over an extended period will appreciably reduce the subsequent degassing rate and indeed some vacuum engineers suggest that baking at a higher temperature is unnecessary. However, most workers recommend baking to 400–450°C for at least 16 hours to obtain ultimate degassing rates which are compatible with ultrahigh vacuum pressure attainment. A number of values have been reported for outgassing rates from metal after such bakeout processes, most of them being concerned with stainless steel. Values vary from 10^{-10} to 10^{-12} Pa $m^3 s^{-1} m^{-2}$ at room temperature and one can reasonably expect to attain the upper limit for most metal systems where the precautions and processes described in this section have been adhered to. It must be remembered, however, that if the system is let up to air, especially if the air is not dried, a re-bake would be required on subsequent pump-down to remove surface adsorbed gas.

2.3.3 Permeation of gases through metal

If the metal forms part of the vacuum envelope, a further criterion arises from the permeation of gas through the metal. Unlike glass, gas diffusion takes place through the crystal lattice and only those gases which are soluble in the metal will permeate through it. Thus helium and the other inert gases will not permeate through any metal even at elevated temperatures, whereas hydrogen and oxygen will permeate to some extent through most metals. Permeation of the diatomic molecular gases varies as the square root of the pressure difference, i.e. n in Equation (2.1) is $\frac{1}{2}$. This indicates that the molecules dissociate on adsorption at the high pressure surface and diffuse through the metal as atoms, recombining on the vacuum side on desorption. Hydrogen, having the highest diffusion rate, is the main gas involved in permeation through metals and in *Figure 2.8* the log of the permeation constant for hydrogen through several metals has been plotted against the reciprocal of the temperature. The values have been derived from curves given in a survey by Norton[19]. Since K for metal system is dependent on $(P)^{1/2}$, it is expressed as $m^2(Pa)^{1/2} s^{-1}$ when P, Q, etc. are in SI units.

It can be seen from the figures that the permeation rate of hydrogen through palladium is about two orders higher than for any other metal. The permeation of other gases through palladium on the other hand is negligible, so that palladium makes a useful filter for obtaining pure hydrogen, usually employed in the form of a heated tube. The permeation rate of hydrogen through nickel and iron is also relatively high. For example, at pressures below 10 Pa it is higher than the permeation rate of helium through silica under similar conditions, particularly at elevated temperatures (cf. *Figure 2.4*). The 'glass-sealing' metals of the fernico type have similar high permeation rates. A review of permeation and outgassing of vacuum materials by Perkins[20] list values of K for various stainless steels and iron–cobalt–nickel alloys.

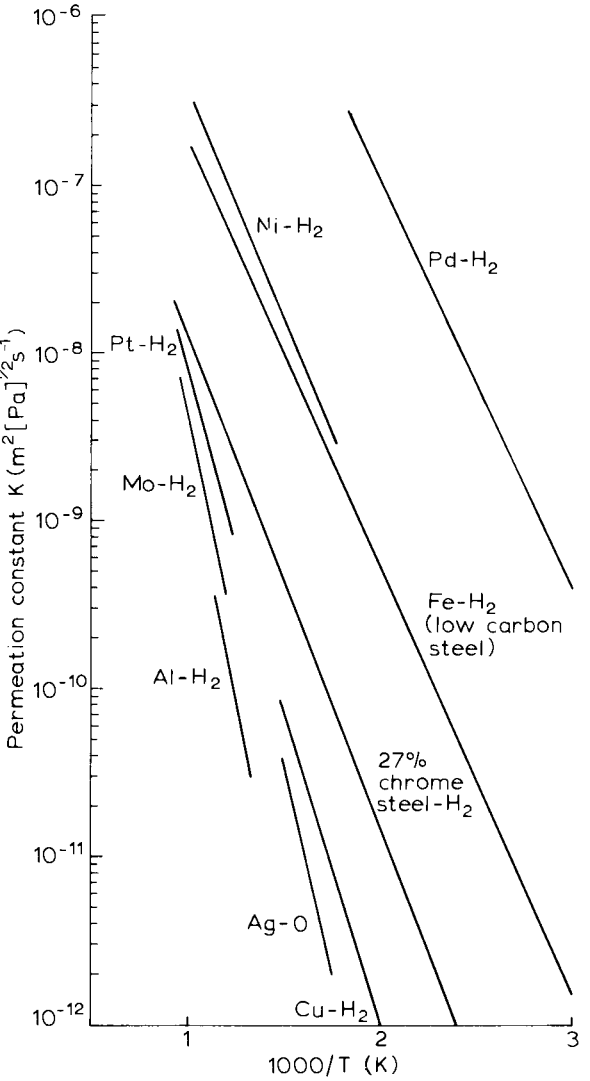


Figure 2.8 Permeation constant for hydrogen through various metals as a function of temperature (from Norton[19])

The permeation constant for stainless steel is some two orders lower than for the alloys, nevertheless hydrogen permeation through chamber walls constructed of stainless steel, due to the partial pressure of hydrogen in the atmosphere, cannot be ignored if ultrahigh vacuum pressures are to be achieved. This can be illustrated by considering the influx of hydrogen from the atmosphere into a vacuum vessel for the same example as used in Section 2.2.2, i.e. a sealed-off spherical vessel of 16 mm radius with a wall thickness of 1 mm. The rise in pressure with time in the vessel made of various metals is plotted in *Figure 2.9*. In spite of the low partial pressure of hydrogen in the atmosphere (5×10^{-2} Pa) the pressure rises, at room temperature, to 10^{-5} Pa

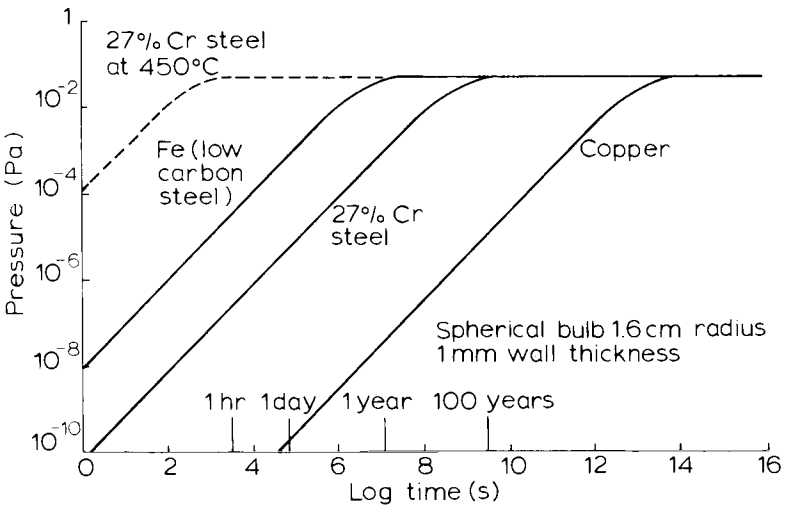


Figure 2.9 Accumulation of hydrogen in a metal vacuum bulb as a result of permeation from the atmosphere at 25°C for various metals

in less than an hour for iron, two or three days for stainless steel and about a hundred years for copper. At higher temperatures the permeation increases markedly and if the stainless steel sphere were to be baked at 450°C then the pressure would rise above 10^{-4} Pa in less than a second (see dotted curve, *Figure 2.9*).

In the case of iron, the influx of hydrogen can also occur as a result of the chemical reaction of water with the exterior walls. The reaction produces hydrogen, some of which can permeate to the interior. Since the abundance of bound hydrogen, as water vapour, in the atmosphere is about 10^5 times that of free hydrogen, rusting is probably a much greater source of hydrogen influx than permeation of the free hydrogen from the air.

The permeation rates of other gases through most metals are at least an order lower than for hydrogen and for practical purposes can be neglected. However, the permeation of oxygen through silver is of interest. Because of the high solubility of oxygen in silver, its permeation rate is much higher than any other gas including hydrogen. As a result, a heated silver tube can be used to admit oxygen to a vacuum system. The oxygen so produced is spectrographically pure.

2.3.4 Physical requirements

When the metal is employed as the vacuum envelope, then there are further criteria imposed on it which concern its physical and chemical properties. Clearly the envelope must be mechanically strong enough to withstand the pressure due to the atmospheric environment and one does not wish to employ metals which would entail excessively thick walls to satisfy this criterion. Also it must maintain its strength when heated to baking temperature and indeed should not be distorted by temperature cycling. The need to fabricate the shape of the vacuum vessel and to provide leak-tight joins by welding and

brazing, further restricts the choice of metals. Another important feature is the chemical stability of the metal. If the metal reacts with the atmospheric gases when heated then erosion of the surface and mechanical weakening may occur. Copper is unsuitable for this reason; oxidation at elevated temperature causes scaling which flakes off to expose fresh surfaces. As has already been pointed out, rusting of mild steel gives rise to appreciable hydrogen permeation, which makes it unsuitable for ultrahigh vacuum systems. Finally, in many applications a vacuum envelope is required which is non-magnetic.

Bearing these requirements in mind, present-day technology favours stainless steel as the most satisfactory metal for ultrahigh vacuum. It is also a fairly common material and reasonably inexpensive. Stainless steel is a term commonly used to indicate any or all iron alloys which resist atmospheric corrosion. It generally refers to low carbon steels containing 10–25% chromium and two main classes are recognized: (1) those containing chromium as the only major alloying constituent and (2) those containing both chromium and nickel as the major alloying elements. The latter group, known as the Austenitic stainless steels, American AISI 300 series, are of most interest, particularly those containing 18% chromium and 8% nickel (18/8 stainless steels). They can be considered as non-magnetic, having magnetic permeabilities of less than 1.02. They resist corrosion up to 800°C and do not harden when hot worked. On the other hand they can develop high mechanical strength when suitably processed and maintain their strength at elevated temperatures. For argon arc welding the steels are either 'stabilized' by the introduction of a small percentage of titanium or niobium, or a specially low carbon content Austenitic steel is used. *Table 2.5* gives the maximum percentage composition of the various stainless steels in common use in the UK together with the American equivalent or near equivalent type numbers. The low carbon content steel is preferred and used extensively in the States, type 304. It is now becoming more used in the UK where previously it has been difficult to obtain. EN58B or EN58F are also employed. All three appear to be perfectly satisfactory for most ultrahigh vacuum systems. EN58F is the most readily available and cheapest but has a higher magnetic permeability making it unsuitable for some applications.

TABLE 2.5. Composition and type numbers of Austenitic stainless steels

British type	*American equivalent AISI No.*	*Composition (maxima) (%)*						*Description*
		C	*Si*	*Mn*	*Ni*	*Cr*	*Others*	
EN58A	320	0.16	0.20 (min.)	2.0	7–10	17–20		18/8 steel
EN58B	321	0.15	0.20 (min.)	2.0	7–10	17–20	Ti 4 × C	18/8 steel Ti stabilized for welding
EN58E	304	0.08	0.20 (min.)	2.0	8–11	17.5–20		Low carbon 18/8
EN58F	347	0.15	0.20	2.0	7–10	17–20	Nb 8 × C	18/8 steel Nb stabilized for welding
–	430	0.12	1.0	1.0	–	14–18		Magnetic

2.4 Ceramics

The term 'ceramic' describes a wide range of inorganic non-metallic compounds which have attained a hard solid crystalline state as a result of firing. They have advantages over glass as a constructional material for insulation application, in their better mechanical strength, particularly at elevated temperature, their better electrical properties and in their capability of being fabricated with close dimensional tolerances. Like glass they are practically chemically inactive and have a low vapour pressure. Advances in the technology and manufacture of ceramics have produced improved quality components at reasonable cost and as a result they are finding increasing use in vacuum systems and devices.

There are three basic types of ceramics: pure oxides, silicates and special types of nitrides, borides and carbides. Those in the last group have been developed specifically for high temperature applications in rockets, etc. and have not found application so far in vacuum technology.

Although the oxides can be single-phase crystalline compounds, most ceramics also contain a certain proportion of glassy phase material which bonds the crystal aggregates together. Since the ceramic is formed by a sintering process it is usually porous in structure, a 10% pore volume being quite common, and the glassy phase also serves to seal the pores and render the ceramic gas tight. In general the silicate ceramics have a higher quantity of glassy phase as, for example, in porcelains where up to 70% can be of this phase. The amount of glassy phase present has a marked influence on the ceramic properties, in particular the mechanical strength and electrical properties. The method of fabrication can also affect the properties giving differing microstructures to the resulting ceramics.

In general the oxide ceramics are obtained synthetically from chemically prepared materials, whereas the silicates are obtained from naturally occurring minerals, although to some extent they may be purified by preparatory processes. Because of this the silicates tend to show more variability and even when they are substantially of the same chemical composition, can have differing properties according to the source of the materials.

Since the vacuum technologist requires high quality materials with controlled characteristics which are reproducible, the oxide ceramics are usually preferred. However, there are certain properties of the silicates, for example their dielectric constant, which makes them desirable for certain applications. Also they are easier to make and, therefore, less expensive.

The manufacture of ceramics depends, to some extent, on the type, but the basic stages in the schedule apply to most ceramics. It involves grinding the ingredients to fine particles and then adding sufficient moisture, or in some cases organic binder, to give the mix a clay-like plasticity. The clay is then worked to de-aerate it and produce a homogeneous mass. The ceramic part is moulded, pressed, or extruded from the homogeneous clay and left to dry. When dry it is sintered in a high temperature furnace to give the final hard state. During this firing process shrinkage of the component takes place, which has to be allowed for. In some cases it is possible to give an intermediate lower temperature firing, which permits the ceramic to be machined before the final

hardening. For more details of the various processes the reader is referred to Espe[21].

2.4.1 Types of ceramics for vacuum application

Only a limited number of ceramic types from the large range available are suitable for vacuum components. If the ceramic is to be used for the vacuum envelope, necessitating ceramic–metal or ceramic–glass seals, then the range is further restricted.

The main impetus to the use of ceramics as insulators in vacuum systems instead of glass, arose from the work carried out in Germany and in the USA during the 1940s on suitable ceramics with low dielectric loss factors for microwave tube components. The ceramics developed were based on steatite, a magnesium metasilicate ($MgSiO_3$) body ceramic, and the main advance was in the successful ceramic-to-metal seals which were evolved. Typically steatite is made by combining 70–80% talc ($3MgO.4SiO_2.H_2O$) with 20–30% china clay ($Al_2O_3.2SiO_2.2H_2O$) to which alkali or alkaline–earth oxides are added as flux. The final ceramic, which is fired around 1400°C, consists of the $MgSiO_3$ crystals bonded by a glass, high in alkali oxides. Steatites have a limited temperature range for firing, ± 10°C, and require accurate temperature control in manufacture. An improved material is obtained by enriching steatite with magnesium compounds to give another type of ceramic known as Forsterite (Mg_2SiO_4).

Forsterite has a wider firing temperature range and also lower dielectric loss. The thermal expansion coefficient is high 11×10^{-7} per °C, i.e. similar to that of soft glasses, and, therefore, it can be sealed to chrome–iron and also titanium to which it is more closely matched. Another ceramic used in the 1940s and early 1950s was zircon porcelain ($ZrO_2.SiO_2$). Zircon porcelain has a very low coefficient of thermal expansion, closely matched to molybdenum, and as a result has a good heat shock resistance. It has, however, a high dielectric constant.

At that time pure oxide ceramics were too expensive for general use, although their superior properties were appreciated. With the increased demand and improved ceramic technology, such ceramics have become more readily available, especially alumina ceramics which are almost extensively used for vacuum applications today. A variety of alumina-body (Al_2O_3) ceramics are available, having an alumina content of from 85% to nearly 100%. They are mechanically stronger than most other ceramics and in spite of their higher thermal expansion coefficient than for zircon porcelain, they are able to withstand high temperatures. The mechanical strength and dielectric properties improve with increase in purity but the purer alumina-body ceramics are more difficult to make and, therefore, more expensive. The purity also affects the ease with which the ceramic can be metallized. The sintering temperature for high alumina-body ceramics is higher than for silicate ceramics, from 1700–1850°C.

Other oxides that have found use in vacuum applications are zircon (ZrO_2) and beryllia (BeO). Beryllia has an advantage over alumina in having a high thermal conductivity but in powder form it represents a serious health hazard and must, therefore, be used with great caution.

An interesting development has been the production of glass ceramics, known as Pyroceram* or Cervit†. By heat treating glass, it has been found possible to convert it to a crystalline ceramic state, particularly if nucleating agents are added to the glass body. The component can be fabricated by normal glass making techniques and converted by further heat treatment. The resultant material is an opaque ceramic of virtually the same dimensions as the original glass article, but with greater strength and resistance to thermal shock.

A further development arising from these materials has been the introduction of a machinable glass ceramic from Corning, trade name Macor[22], which can be machined in the final state with standard metal-working tools. The parent glass is a heavily phase separated white opal glass containing fluorine-rich droplets. On subsequent heating to 825°C, plate-like crystals of mica phase fluorophlogopite ($KMg_3AlSi_3O_{10}F_2$) are formed. The result is a microstructure consisting of a highly interlocked array of two-dimensional mica crystals dispersed in a brittle glassy matrix. During machining the fracture is localized by deflection, branching and blunting of the cracks by the mica crystals and the enclosed glassy regions are dislodged. The machining resolution is proportional to the size of the mica crystals which are typically of the order of 20 μm diameter.

2.4.2 Physical properties

As with glass, the important physical property of a ceramic in its vacuum application is its mechanical strength and its behaviour with temperature, especially where ceramic–metal or ceramic–glass seals are involved. Like glass, ceramics are brittle in that they fracture under strain without any elongation or flow. Also, it is the tensile stress which is of most concern, since the compressive strength is some ten to twenty times higher than the tensile strength. For most silicate ceramics the tensile strength is similar to that of glass, 4 to 10×10^6 kg m^{-2}, but for the alumina-body ceramics, tensile strengths up to 26×10^6 kg m^{-2} can be attained; typical values are given in *Table 2.6*, where the characteristics of the common ceramics used in vacuum are listed. The strength depends on the porosity of the ceramic and also the shape and total cross-sectional area; small diameter fibres are stronger than larger diameter rods.

Changes in temperature set up stresses at the interface with other materials due to differences in the expansion coefficients. However, because of the higher strength, especially under compression, it is possible to make successful metal–ceramic seals without the expansion coefficients of the two materials being as closely matched as is required for glass. If they are closely matched, a more versatile and reliable seal can be obtained than is possible with glass. Typical thermal expansion curves are given in *Figure 2.10*. Unlike glass the curves are almost linear up to the softening temperature. Expansion coefficients of the ceramic types are also listed in *Table 2.6*. Generally the resistance of a ceramic to thermal shock is greater the smaller the expansion coefficient, although the tensile strength is also of importance. From this point of view they are more resistive to thermal shock than glass.

* Tradename of Corning Glass Works
† Tradename of Owen-Illinois Glass

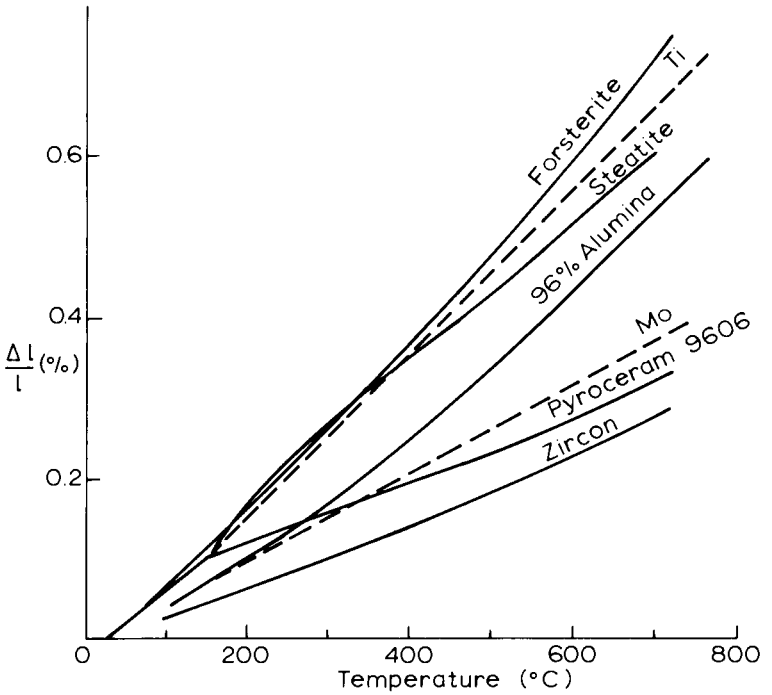


Figure 2.10 Thermal expansion characteristics of some of the ceramics commonly used in vacuum technology and of metals to which they may be sealed

TABLE 2.6. Physical properties of ceramics used in vacuum technology

Ceramic	*Main body composition*	*Coefficient of expansion* ($\times 10^{-7}$ °C^{-1})	*Softening temperature* (°C)	*Mechanical strength/ tensile strength* (kg m^{-2})	*Specific gravity*
Steatite	$MgOSiO_2$	70–90	1400	6×10^6	2.6
Forsterite	$2MgOSiO_2$	90–120	1400	7×10^6	2.9
Zircon porcelain	ZnO_2SiO_2	30–50	1500	8×10^6	3.7
85% alumina	Al_2O_3	50–70	1400	14×10^6	3.4
95% alumina	Al_2O_3	50–70	1650	18×10^6	3.6
98% alumina	Al_2O_3	50–80	1700	20×10^6	3.8
Pyroceram 9606		57	1250	14×10^6	
Macor		94	>1000		2.5

Ceramics do not exhibit the 'slow' monotonic changes in viscosity with temperature experienced with glass. Nevertheless, because of the glassy phase the melting point of a ceramic is often rather indeterminate. A softening temperature is usually defined by the collapse of a cone of given dimensions made from the ceramic. Typical values of softening temperature are given in *Table 2.6*. As a practical guide, ceramics used as part of the vacuum envelope should not be heated in excess of a temperature which is 400–500°C below the

softening temperature. Since the softening temperatures of the ceramics in common use in vacuum systems are above 1200°C, this limitation does not normally present a problem to the vacuum engineer.

2.4.3 Permeation of gases through ceramics

Gas permeates through ceramics in the same way as it permeates through glass, via the pores in the micro-structure. Consequently the permeation rate of gas through ceramic will depend on the packing density of the ceramic 'crystals' (porosity) and whether or not a glassy phase is present. It will also depend on the size of the gas molecule involved, helium having the highest permeation.

The manufacturing process, size of particles, etc., as well as chemical composition, will affect the permeation; for example some ceramics are especially designed to be porous and can be used for leaking gas into a vacuum system at a controlled rate.

In spite of the increasing use of ceramic as the vacuum envelope in tubes and systems, very little information has been published on helium permeation. Most experiments have shown helium leak rates which are lower than for glass without giving details of the permeation constant. Manufacturers claim their material to be vacuum tight, using a helium leak detector with sample discs. There is, however, one article by Miller and Shepard[23] where the permeation of helium and air through Pyroceram 9606 is compared to that through a 97% alumina ceramic. The curves are reproduced in *Figure 2.11*. Also included in the figure is the curve for Macor taken from an article by Altemose and Kacyon[24]. Permeation of oxygen, nitrogen and argon through extruded

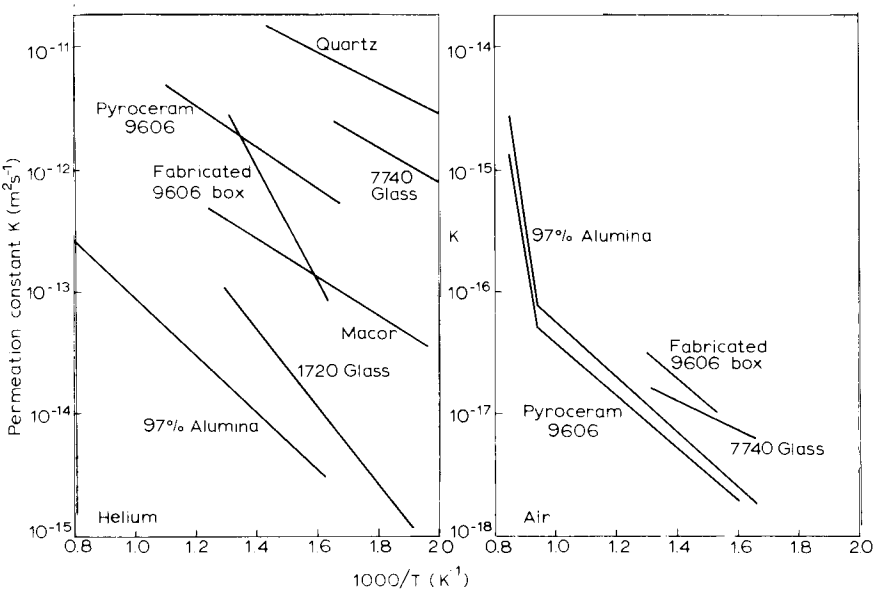


Figure 2.11 Permeation constant for helium through Pyroceram, Alumina (from Miller and Shepard[23]) and Macor (from Corning data sheet)

alumina ceramic tubes at elevated temperatures have been measured by Hayes *et al.*[25]. Below 1500°C the ceramics were impervious within the limits of measurement 1.5×10^{-13} m^2s^{-1}. At temperatures above 1500°C, where permeation constants of 10^{-11} m^2s^{-1} were measured, changes in the microstructure of the tubes took place in some of the materials, giving rise to selective permeation of oxygen. Much higher values were obtained with hot pressed discs of similar composition[26] suggesting that the method of manufacture was the important factor.

In general, provided that the ceramic is designed for vacuum application and has no faults such as cracks or holes, helium permeation does not appear to be a problem in the attainment of ultrahigh vacuum. The situation is further improved since much greater wall thicknesses are possible in ceramic systems without danger of thermal stresses causing fracture.

2.4.4 Outgassing of ceramics

Because of their glassy phase, silicate ceramics behave similarly to glass with regard to gas desorption. Water vapour, carbon dioxide and carbon monoxide are the main gases evolved, the gases coming from adsorbed surface layers and from the bulk of the material. Hydrogen is also usually detected and Gibbons[27] stated that it was the principal gas given off from alumina ceramic chambers designed for a synchrotron storage ring. Clearly the porosity of the ceramic will affect the outgassing properties. For internal components, the degassing is facilitated by using ceramics with high porosity, the converse to the permeation situation. In general the silicate ceramics with a high percentage of glassy phase, such as porcelain, contain the most adsorbed gas, comparable to glass. Steatite, Forsterite and alumina-body ceramic have a much lower gas content and can be further improved by forming the component in a vacuum environment. The advantage of ceramics over glass is the possibility of degassing the component prior to assembly, in the same way that metal parts are pre-stoved. This is probably only necessary where the ceramic part forms an appreciable part of the final vacuum system. In the final vacuum system bakeout temperatures above 500°C can be employed. However, Norton[28] has pointed out that high baking temperatures can cause evolution of gas if metal oxide impurities are present due to dissociation. For example Fe_2O_3, which is a common impurity in silicate ceramics, has a dissociation pressure of about 10^{-4} Pa at 800°C, and can cause prolonged evolution of oxygen when baked at elevated temperatures.

In general, therefore, a ceramic component should be vacuum stoved at about 1000°C before assembly, and baked at 450°C or above for several hours in the final system. The resulting outgassing at room temperature should then be of the order of 10^{-11} Pa $m^3s^{-1}m^{-2}$ or better.

2.5 Other materials

In recent years a number of synthetic materials—plastics, elastomers, epoxy resins, etc.—have found use in high vacuum applications. These materials all exhibit high outgassing rates from gases absorbed during manufacture and also are permeable to some extent to all gases. The values of degassing rate and

permeability are considerably higher than for metal, glass and ceramic. Although the outgassing rate can be reduced by prolonged pumping, or in some cases a low-temperature bake, the materials rapidly re-absorb gas when exposed to the atmosphere, especially water vapour, and values comparable to a baked metal or glass system cannot be achieved. Such materials are, therefore, generally unsuitable for ultrahigh vacuum applications. However, because they are cheap and have chemical and physical properties which are particularly desirable, for example their high elasticity, methods of employing them to give the minimum gas influx have been advocated, particularly in continuously pumped systems where pressures much below 10^{-6} Pa are not required.

Several workers have investigated degassing and permeation characteristics of such materials[29,30,31,32], and data on a wide range of commercially available plastics and elastomers have been published.

In *Table 2.7*, some typical values of outgassing rates at room temperature are given for materials of most interest to the ultrahigh vacuum engineer. The data are taken mainly from the work of Barton and Govier[33] but some other results are also included. They show the effect of baking and exposure to air. For all the materials given, the outgassing rate initially is high 10^{-3} to 10^{-4} Pa $m^3 s^{-1} m^{-2}$ and even after pumping for 50 hours it is not reduced much below 10^{-5} Pa $m^3 s^{-1} m^{-2}$. The exception appears to be PTFE according to Barton and Govier, where values around 10^{-7} Pa $m^3 s^{-1} m^{-2}$ were recorded. This does not agree with finding by Markley *et al.*[37] who reported a reduction of only an order after pumping for 300–400 hours. A large percentage of the gas given off by the synthetic materials is water vapour and baking at the maximum permissible temperature, restricted by deformation or decomposition, reduces the outgassing rate by at least an order. The materials giving the lowest outgassing values in general are those which can be baked to the highest temperature. Thus PTFE, Viton A, Mycalex and Polyimide have all found uses in vacuum systems pumping down below 10^{-6} Pa. Two rather more recent materials introduced by Du Pont namely Kalrez and Viton E60C offer lower outgassing characteristics but have not as yet been fully exploited.

PTFE (Polytetrafluoroethylene) is a fairly hard material, used extensively in the domestic market for non-stick coatings, which is chemically very stable and can be heated up to 300°C. Its important attribute is the low coefficient of friction of its surface against other materials, which allows it to be used as bush bearings for moving parts without the need for lubrication. This accounts for its main application in vacuum engineering. Since fairly small quantities of material are required for bush bearings, the outgassing rate can be tolerated, especially for pressures in the region of 10^{-6} Pa. It can also be used as a coating for hard metal gaskets in demountable seals (cf. Section 6.2). Viton A is a fluorocarbon elastomer suitable for 'O'-ring seals, which can be baked up to 200°C. It replaces neoprene in high vacuum systems with ultimate pressures going down to 10^{-7} Pa. Micalex is not a true plastic being synthetic mica bonded by glass. Its asset is that it can be fairly readily machined and might be considered as the forerunner to the machinable glass–ceramic, Macor.

Polyimide is probably the most interesting material since its outgassing rate after baking to 300°C is lower than most other synthetic materials. However, polyimide has a similar chemical composition to nylon and is to some extent hygroscopic, so that on exposure to air, water vapour is readily adsorbed. The

TABLE 2.7. Outgassing rate for synthetic materials

Reference	*Material*	*Before bake*	*Hours pumping Inc 24 h bake*	*After exposure to air for 24 h*	*Baking temperature* (°C)	*Outgassing rate at room temperature* ($Pa\ m^3 s^{-1} m^{-2}$)
33	Araldite ATI	51				3.4×10^{-4}
			80		85	Not detected
				51		6.0×10^{-5}
33	Mycalex	52				2.7×10^{-6}
			83		300	Not detected
				51		1.0×10^{-6}
33	Nylon 31	51				1.1×10^{-4}
			82		120	8.0×10^{-7}
				51		8.9×10^{-6}
33	Perspex	51				1.3×10^{-4}
			102		85	7.8×10^{-6}
				51		6.6×10^{-5}
33	Polythene	262				4.0×10^{-5}
			496		80	6.6×10^{-6}
				94		2.3×10^{-6}
34	PTFE	10	–	–	–	2.0×10^{-4}
33		48	–	–	–	4.7×10^{-7}
33	Viton A	51				1.3×10^{-4}
			101		200	2.7×10^{-6}
				48		4.0×10^{-5}
35	Polyimide		12 h bake		200	6.6×10^{-8}
35			12 h bake		300	4.0×10^{-8}
				5 h exposure to air		5.3×10^{-5}
36	Kalrez		Unspecified		300	4.0×10^{-8}
36	Viton E60C		Unspecified		300	3.0×10^{-8} ($\sim 10^{-6}$ at 150°C)

unbaked degassing rate can therefore be fairly high. It is a fairly hard resilient material which can be used as a gasket. It has a high expansion coefficient and, since it will suffer permanent distortion if compressed more than 20%, rather more care has to be taken in the design of seals where it is to be used. It can also be used for sealing small leaks in the vacuum envelope. For this purpose the polyimide is dissolved in a suitable solvent to form a viscous liquid which is applied to the leak by brushing. Other leak sealants claimed to be suitable for small leaks in ultrahigh vacuum systems are Araldite, silicone resins and anaerobic polymers. The anaerobic polymers are dimethacrylate ester liquids which polymerize in the absence of oxygen, i.e. when they come in contact with the vacuum. Their vacuum performance is reported by Kendall[38]. Kalrez is the trade name for a perfluoroelastomer which is suitable for 'O'-ring seals and which can be baked up to 300°C. De Chernatony[36] reports that after baking the outgassing rate is similar to that of polyimide. Viton E60C is a variant on the Viton fluoroelastomers which has similar vacuum properties to Kalrez although its outgassing rate at elevated temperatures is not as good.

The values of the permeation constant for the gases hydrogen, helium, oxygen and nitrogen, passing through some of the synthetic materials useful for vacuum work, are listed in *Table 2.8*. The data were taken mainly from Bailey[39] and are from measurements made by Barton at AERE. They are in general agreement with results quoted by other workers, although it can be expected that variations will exist on materials from different manufacturers.

Since the diffusion of gas is via the pores in the structure, the permeation is proportional to pressure and the values can be compared directly with those for glass or silica. The values for helium permeation through synthetic materials are at least an order greater than for fused silica and it will be noticed that the hydrogen permeation rate does not differ greatly from that of helium. Indeed there is not the marked effect of molecular diameter found with glasses, so that nitrogen permeation from the atmosphere is as serious as helium permeation. The high value for PTFE is associated with the difficulty of manufacturing the material with a higher (less porous) density. As far as is known there are no published values for the permeation of gas through Kalrez or Viton E20C and the only values of gas permeability through polyimide are those of George[41] quoted by Perkins[20].

The main use of plastic materials in high and ultrahigh vacuum is as gaskets in demountable seals and valves. In such applications a minimum surface area is exposed to the vacuum and the diffusion path through the material is relatively large. This is discussed in Chapter 6. A useful review on the selection of elastomer materials for seals has been given by Peacock[42].

Before leaving this section on synthetic vacuum material, mention should be made of mica, which is used extensively in vacuum devices, such as electron tubes. Mica is a naturally occurring mineral, which is in the form of transparent laminations. It is either K-Al-silicate or K-Mg-Al-double silicate and it can also be made synthetically in commercial quantities. It is a good insulating material, resistivity 10^{16} to $10^{17}\,\Omega.\mathrm{cm}$, with a high dielectric constant. One of its special properties is the ease with which it may be split into thin sheets, which can be stamped out into relatively complex shapes. It is, therefore, particularly suitable as spacers between electrodes, which need to be accurately located.

Unfortunately the laminated structure of mica results in gas being trapped

TABLE 2.8. Permeability constants at ambient temperature of synthetic materials

Material	*Permeability constant K in m^2s^{-1} at 23°C*				
	Nitrogen	*Oxygen*	*Hydrogen*	*Helium*	*Argon*
Polythene	9.9×10^{-13}	3.0×10^{-12}	8.2×10^{-12}	5.7×10^{-12}	2.7×10^{-12}
PTFE	2.5×10^{-12}	8.2×10^{-12}	2.0×10^{-11}	5.7×10^{-10}	4.8×10^{-12}
Perspex	–	–	2.7×10^{-12}	5.7×10^{-12}	–
Nylon 31	–	–	1.3×10^{-13}	3.0×10^{-13}	–
Polystyrene	–	5.1×10^{-13}	1.3×10^{-11}	1.3×10^{-11}	–
Polystyrene*	6.4×10^{-12}	2.0×10^{-11}	7.4×10^{-11}	–	–
Polyethelene*	$6–11 \times 10^{-13}$	$2.5–3.4 \times 10^{-12}$	$6–12 \times 10^{-12}$	$4–5.7 \times 10^{-12}$	–
Mylar 25-V-200*	–	–	4.8×10^{-13}	8.0×10^{-13}	–
CS2368B (Neoprene)	2.1×10^{-13}	1.5×10^{-12}	8.2×10^{-12}	7.9×10^{-12}	1.3×10^{-12}
Viton A	–	–	2.2×10^{-12}	8.2×10^{-12}	–
Kapton** (polyimide)	3.2×10^{-14}	1.1×10^{-13}	1.2×10^{-12}	2.1×10^{-12}	–

* Data taken from work of Brubaker and Kammermeyer[40]
** Data taken from the work of George[41]

between the layers, which is difficult to drive off. Baking at too high a temperature drives off the water of crystallization and causes the mica to crumble. The only satisfactory method of outgassing it is prolonged baking at say 200–300°C. Synthetic mica is better from this point of view, but on the whole mica is not a very satisfactory material for ultrahigh vacuum applications.

The permeability of gas, normal to the cleavage plane is low, at 400°C, K for helium is less than 10^{-7} m^2s^{-1}, i.e. it is lower than for the best glass. This factor, coupled with its strength in the same direction, has promoted its use as vacuum windows, which can be thin and therefore transmit a high percentage of incident radiation.

The thermal expansion coefficient of mica is high, 80 to 130×10^{-7} $°C^{-1}$ parallel to the cleavage plane and 160 to 250×10^{-7} $°C^{-1}$ perpendicular to it. It is most nearly matched to the soft glasses and metals with similar expansion coefficients such as nickel–iron or chrome–iron.

2.6 Fabricating techniques

The method of fabricating glass vacuum systems depends on fairly standard glass working techniques, as for instance those employed in the manufacture of electronic tubes. For details of such techniques the reader is referred to the comprehensive book by Espe[21], which covers the construction of the components, bulbs, feet, etc. and the joining of the components to form completed devices. Similarly the manufacture of metal piece parts follows standard engineering techniques of machining, forging, etc., although there are certain precautions which have to be taken in the design to facilitate outgassing. This also applies to the manufacture of ceramic components. It is beyond the scope of this book to describe the general techniques of fabricating the components or piece parts of an ultrahigh vacuum system and this section will deal only with the sealing of the parts together and on special conditions dictated by the vacuum requirements. The seals referred to are the permanent joins in the vacuum system, either between similar materials or dissimilar materials. Demountable seals are dealt with in Chapter 6.

2.6.1 Glass-to-glass seals

Apart from one or two exceptions, glass is sealed to itself and to other materials by a fusion process, whereby the temperature is raised well above the softening point of the glass. Since glass is a brittle material the parameter of prime importance in making a successful seal is the thermal expansion coefficient of the components which must be reasonably matched over a wide temperature range. Two glasses of different composition can usually be successfully sealed together if the average thermal expansion coefficients of the two glasses do not differ by more than 10% and if their transition temperatures are similar. Also both glasses must dissolve well in each other.

If two glasses cannot be joined because of their thermal expansion differences, then intermediate glasses with expansion coefficients lying between those of the two glasses can be employed to make what is termed a graded seal. In going from soft glass to Pyrex, several intermediate glasses

(about 7) are required to maintain the 10% criterion at each join and considerable glass blowing skill is called upon to maintain an even wall thickness. The length, l, of the individual sections of a graded seal must be large enough for the stresses at the ends to be sufficiently attenuated. Lewin and Mark[43] have looked at this problem theoretically and for glass tubes have given a criterion which is safe in practice of

$$l \geqslant 0.85(ah)^{1/2}$$

where a is the tube radius and h the wall thickness. To minimize the stresses it is important that any join in glass especially graded seals should be carefully annealed, preferably in an oven. To anneal the glass it should be held above the annealing temperature for as long as possible, up to an hour for large components, and then cooled very slowly to room temperature.

An alternative method of sealing glass parts of similar expansion coefficients is to use a glass solder. These are glasses of low softening point, significantly lower than the softening point of the parts to be joined. In this way the glass parts are not distorted during sealing. For example for sealing soda glass parts together a solder glass that will soften at 300–400°C is used. The technique is particularly useful for sealing flat windows or lenses. Usually the solder glass is applied, in the form of fine granules suspended in a nitrocellulose binder, between the components to be joined by brush or a similar coating method. It can also be pressed as a sintered preform, for example a ring, or applied as a fused casting. The components are then heated in an oven to the soldering temperature and the parts pressed together, possibly under their own weight. The expansion coefficient of the solder glass must also closely match that of the components and in this respect solder glasses for soft glass are more readily available than those for hard glasses. In general the solder glasses are borate based, and do not have such good chemical resistance as silicate glass. On the other hand their chemical activity gives them good bonding to the silicate glasses. 'Pyroceram' cements are also available for sealing glass parts together. They have the advantage over solder glasses that after melting they devitrify to become a ceramic-like material with a much higher softening point. Such seals can therefore be baked to higher temperatures for outgassing purposes.

2.6.2 Glass-to-metal seals

Glass-to-metal seals are rather more complex and the degree of matching of the expansion coefficients required depends on the shape of the seal, the plasticity of the metal and the annealing procedure. There is also the problem of obtaining a good vacuum-tight bond between the metal and glass. For this one normally relies on a layer of metal oxide which, to some extent, will dissolve in the glass to give good adherence. For some metal–glass combinations the oxide layer is formed in the sealing process but for others it is necessary to pre-oxidize the metal before sealing. Another requisite for a good seal is that the metal should be thoroughly degassed, otherwise bubbles will appear in the glass at the seal which can lead to gas leakage.

When considering a 'matched' seal, one where the coefficients of expansion of the metal and glass are similar, the characteristic curve of thermal expansion against temperature must be considered over the whole range, from room temperature to the glass softening temperature. As shown in *Figure 2.2*, the

characteristic is nearly straight for glass up to the annealing temperature where it increases markedly. For pure metals the expansion characteristics are almost linear over the same range but do not show such an increase at the higher temperature. There are, therefore, few pure metals which can be sealed satisfactorily to glass without introducing considerable strain. The exceptions are tungsten and molybdenum which can be sealed to borosilicate glasses specially developed for such seals. Tungsten, with an expansion coefficient of $44 \times 10^{-7}°C^{-1}$ can be sealed to 7720 and equivalent glasses, whilst molybdenum with an expansion coefficient of $55 \times 10^{-7}°C^{-1}$ can be sealed to the lower melting point borosilicate glasses such as 7052 (see *Table 2.1*). Such seals are usually restricted to metal wires or pins where the glass is under compressional strain. Care must be taken not to over-oxidize the metal, as both metals readily oxidize to give layers which are not very strongly adherent. The seals formed, however, are reliable for ultrahigh vacuum purposes and particularly useful where non-magnetic seals are essential.

The main 'matched' metal-to-glass seals use alloys of iron specifically developed for glass sealing purposes. The principal alloys used are basically nickel-iron where the expansion coefficient can be varied continuously over a range encompassing most of the industrial glasses by changing the nickel content from 35–60%. The expansion coefficients of these alloys show a rise at the magnetic transition temperature (Curie point) which matches rather well the behaviour of glass at the annealing temperature. Typical curves of expansion against temperature are shown in *Figure 2.12* for the sealing alloys from a British manufacturer. The expansion coefficient of the matching glasses for two of the alloys are also shown in the figure. For the simple nickel–iron alloy, lowering of the nickel content lowers the expansion coefficient but unfortunately also lowers the Curie point. For the Curie point to be above 400°C the nickel content must be above 44% which gives an expansion coefficient over $70 \times 10^{-7}°C^{-1}$. This limits its use to seals with soft glasses. For example a 50–50% Ni–Fe alloy is well matched to 0120 glass, having an average expansion coefficient of $90 \times 10^{-7}°C^{-1}$ and a Curie point of about 500°C (*Figure 2.12*).

To improve the vacuum tightness of the seal a small amount of chromium is often added (0.8–6%). The chromic oxide formed during the manufacture of the seal dissolves well in the glass and is very adherent to the metal. Alternatively the alloy can be plated with a thin copper layer. A seal of this type, known as the Dumet seal, is extensively used in the electron tube industry. The basic Ni–Fe alloy (42% Ni) in the form of a rod is clad by a brazing technique with copper and then drawn. After drawing, the wire is coated with a layer of fused borax (borated), to form red cuprous oxide during the sealing. Cuprous oxide gives a good bond to both the glass and the copper. Although the expansion coefficient of the Ni–Fe alloy is rather lower than most soft glasses, the excellent bond and ductile nature of the copper layer makes it possible to obtain good seals even to lead glass.

The addition of cobalt to nickel–iron, or the partial substitution of the Ni with Co, has the effect of raising the Curie point without materially affecting the expansion coefficient (cf. *Figure 2.12*). It is this system of alloys which is used for sealing to hard glass and is of most interest to the vacuum engineer. The first of these glass-sealing alloys was produced under the trade name of 'Kovar' and consisted of 54% Fe, 29% Ni and 17% Co. This has an average

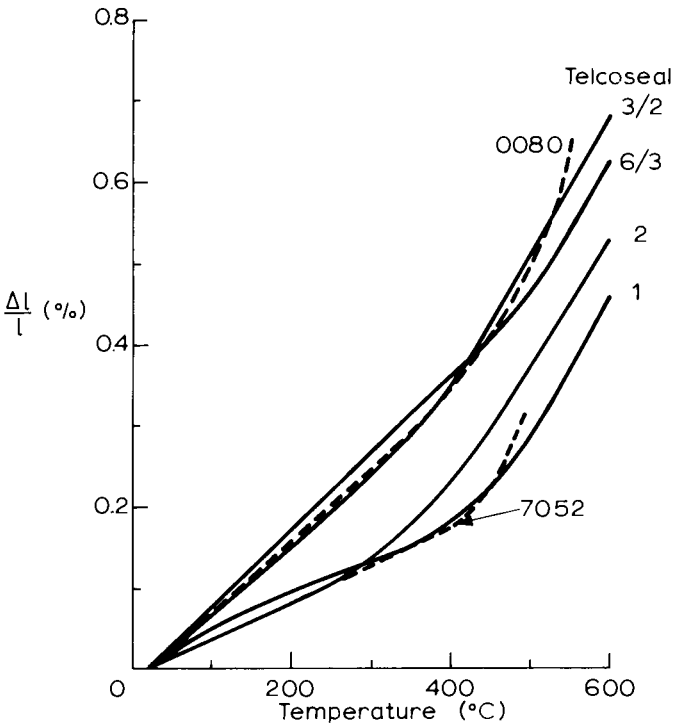


Figure 2.12 Expansion–temperature curves for some glass sealing alloys from Telcon Metals Ltd.

Alloy	*Composition*				*Matching glass*	
	Fe	Ni	Co	Cu	(*Corning*)	(*Schott*)
Telcoseal 1	54	29	17	–	7052	8250
Telcoseal 2	58	42	–	–	7556	–
Telcoseal 3/2	49	47	–	5	0080	8196
Telcoseal 6/3	50	50	–	–	0120	8095

expansion coefficient of $50 \times 10^{-7}°C^{-1}$ and a Curie point of 430°C. It is well matched to the borosilicate glasses of the type 7052. Since the patents have expired, similar alloys are now manufactured under various trade names, such as Vacon 12, Telcoseal I and Nilo K. The material is normally vacuum melted and of high purity and the parts are vacuum stoved before sealing. It is common practice to pre-oxidize before sealing to the glass, although with some glasses, particularly sintered preforms, it is not necessary. Care must be taken in using Kovar at low temperatures because below ~200 K it can change its structure and give a higher expansion coefficient. Special batches for low temperature application are available.

If a Kovar tube is sealed to a glass tube at one end whilst the other is welded to stainless steel, then Lewin and Mark[43] suggest a minimum unglassed length, l, to prevent strain at the glass surface, of

$$l = 3.5(ah)^{1/2}$$

where a is the metal tube radius and h the wall thickness. In any high vacuum system the amount of Kovar forming the vacuum envelope should be kept down as low as possible because of the relatively high hydrogen permeation rate (see Section 2.3.3, *Figure 2.8*).

An alternative alloy for sealing to soft glass is chrome-iron, having a chromium content of 23–28%. It is used extensively in the electronic tube manufacturing industry, partly because it cannot be over-oxidized and thus gives the manufacturer considerable tolerance on the sealing process and partly because of its hardness which makes it possible to plug the lead wires directly into a connecting socket.

Non-matched seals can be made to the ductile metals such as platinum and copper, the strain being relieved by deformation of the metal. An example of this type of seal is the joining of glass to copper tubing with a feathered edge to form the well known 'Housekeeper seal'. Similar seals can be made directly to stainless steel[44]. Alternatively a soft metal gasket of indium or gold can be interposed between the metal and the glass and the seal made with a thermo-compression bond. The temperature used is below the softening point of the glass and metal. Being easily deformed, the indium or gold will take up a certain amount of strain from the mismatch in the thermal expansion coefficients. An extension of this method has been reported where seals can be made between glass and harder metals such as Kovar, at temperatures below the softening point, by applying an electrostatic force between the components[45]. For such a seal the surfaces must be very flat and clean with a smooth polished finish.

2.6.3 Metal seals

There are basically two methods of sealing metal components together to form permanent joins. In one, the contact surfaces are fused together either by melting or by pressure, without any intermediary metal, and in the other, a softer and lower temperature melting point metal is interposed between the two surfaces to be joined and melting of the component surfaces themselves does not take place. The processes are covered loosely by the terms welding for the former class of seal and brazing or soldering for the latter. The terms are not precisely defined since some welding processes involve an intermediary metal and there is no real dividing line between brazing and soldering, the two terms are often applied to the same process. Welding is appropriate when the metals to be joined are similar or readily form alloys, whilst brazing and soldering allow dissimilar metals to be joined. Unlike glass the thermal expansion coefficient of the metal is not so important since metals are more ductile. Metals with quite large differences in heat expansion properties can be satisfactorily sealed by suitably designing the seal to take up the strain, for example by thinning the metal at the join. On the other hand the alloying process can cause embrittlement of the metal at the join which cannot be annealed out.

The process which is chosen will depend not only on the metals to be joined but also on the shape of the components to be joined and the function of the seal. Thus for internal vacuum components the strength of the seal is probably the only criterion, whereas for the vacuum envelope the leak-tightness of the seal is the main consideration. This section will primarily be concerned with

the latter type of seal for metal vacuum system construction where special techniques and precautions have to be employed.

Welding can be carried out by a number of techniques: (1) oxy-acetylene welding; (2) electrical resistance welding; (3) electric arc welding; (4) electron beam welding; (5) laser welding and (6) cold pressure welding.

Oxy-acetylene welding, whereby the metal parts are fused together in the flame of an oxy-acetylene torch, is a widely used process for joining base metals. The resulting weld tends to be porous as a result of the gas taken up by the molten metal and also severe oxidation can take place. For these reasons the technique is inappropriate for high vacuum components.

Electrical resistance welding is generally restricted to limited area welds. The components to be joined are pressed together between copper electrodes and a heavy current passed. Because of the resistance at the point of contact, the current joule heats the metal causing local melting to form a fusion weld. The technique is referred to as spot welding and is extensively used in the fabrication of the electrode structures in electronic vacuum tubes. Spot welding machines have been perfected over the years to give pulse currents of varying duration and amplitude with adjustment of the contact pressure. By suitably selecting these parameters a wide range of metals of differing shapes can be welded together, e.g. tungsten wires on to nickel sheet. The method allows precision welding of quite delicate components and is most suited to this type of application, although it is applied also to heavy engineering such as welding of car bodies for example. The welds are strong, free from contamination and, since they take place over a limited area, do not present a serious source of gas influx in a vacuum system.

An extension of the technique to produce a series of spot welds in close proximity, known as stitch welding, can be used to produce a continuous vacuum-tight weld between overlapping sheet metal components. It can only be applied to relatively thin metal sheets, up to 2 mm thick, but nevertheless has found use in disc seals on the vacuum envelope of electronic tubes where the very localized heating allows the seal to be made close to glass–metal seals for example. Nickel, iron and their alloys including stainless steel are readily spot welded and give good strong joins. The higher melting point metals such as tungsten, molybdenum and tantalum are less satisfactory, unless an intermediary metal is used such as platinum. The highly conducting metals such as silver and copper are also difficult to weld because of their low contact resistance. Aluminium is also difficult to spot weld because of the insulating oxide layer which forms on the surface. Details of the spot welding technique and applications can be found in Espe[12].

Of more interest to the vacuum engineer is the electric arc welding method. Traditionally electric arc welding was applied in air, often with one electrode forming a welding rod which was continuously fed as it melted on to the workpiece by the heat of the arc. It gave an alternative method to oxy-acetylene welding for heavy engineering applications. The welds, however, suffer from the same problems, from the vacuum engineer's point of view, of porosity and oxidation.

Alternatively, the arc can be struck in a protective atmosphere of hydrogen or an inert gas and it is this technique which has become widely used for fabricating stainless steel vacuum systems. Argon gas is used with a torch having a central tungsten electrode with a ceramic tubular shield through

which the gas is passed. For steel the electrode is made negative and the workpiece positive, whilst an a.c. arc is used for aluminium. Various sized torches can be obtained to cover a range of applications, from small piece parts to large vacuum chambers. The small torches taking up to 100 amps are usually air cooled and can be used for internal seals in tubes, etc. The larger torches taking two or three times the current have to be water cooled. Because of the protective atmosphere, no flux is required and fusion takes place over a restricted area so that quite fine fusion welds can be made. Fillet seals can also be made by using a welding rod of the same material as the piece parts. The size of the torch, gas flow, rate of traversing the workpiece, etc., can be varied and are selected for the particular job. The manufacturers of the welding equipment give details of the procedures. Because only a limited area is heated, the components may have to be annealed after welding to relieve strain and distortion can take place which necessitates 'truing up' by machining. Although the weld does not oxidize, the metal just outside the fusion weld which is not protected by the stream of gas can be slightly affected. For easily oxidized metals such as molybdenum and tantalum the welding has to be carried out in an argon-filled box.

For vacuum application care has to be taken in designing the welded join. It is important to avoid trapped volumes which could become virtual leaks in the vacuum system. Similarly it is bad practice to leave cavities in which dirt can accumulate on the vacuum side. *Figure 2.13* gives some examples of good and bad welding practices for various types of joins. In general, wherever possible, the continuous weld should be on the vacuum side and any secondary welds needed for strength should be intermittent and on the atmospheric side of the vacuum system.

Electron beam welding is a fairly recent innovation where the join is heated to the fusion temperature in vacuum by a focused high energy electron beam (>10 kV). Limitation in the size of the vacuum chamber used in such welding machines restricts their application to relatively small piece parts. The intense local heating that can be obtained from the electron beam makes it particularly useful for welding high melting point metals especially if the metals are normally readily oxidized at these temperatures. Thus tungsten-to-tungsten fused welds can be made in such welding machines. It is not of great interest for the fabrication of vacuum envelopes and it is a rather expensive process. A similar method is laser welding, using the energy of a laser beam to heat up the metal. It has the advantage over electron beam welding in not requiring a vacuum environment.

Pressure bonding of metals either at room or elevated temperatures is gaining popularity as a method of forming vacuum type joins. It is applicable to the softer metals such as copper and silver, but can be used with harder metals by interposing a softer metal between the surfaces. An example of the latter is to use indium between Kovar surfaces in a disc seal. The surfaces to be joined have to be free of oxide and the pressure requirements are fairly high $>10\,\mathrm{Kg\,mm^{-2}}$, although they decrease with increasing temperature. The main interest for this type of welding is for metal envelopes of electronic tubes where seal-off stems, attachment of windows, etc. can be carried out without heating up the components. The large forces required do not commend the method for large vacuum systems especially as it is limited to flat surfaces.

For the assembly of some components, for example, where re-entrants are

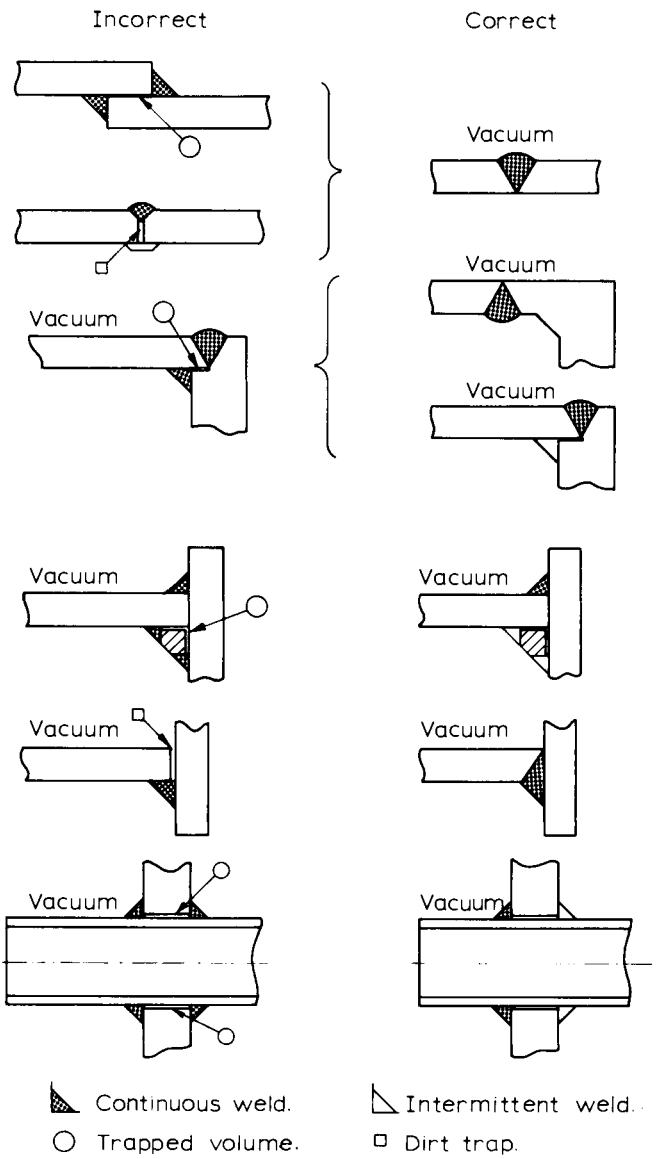


Figure 2.13 Examples of good and bad welding practice according to Kronberger[46]

involved, it is not possible to obtain access to the mating surfaces for welding. It is here that brazing or soldering techniques are used. Low-temperature solders are too limiting on service temperature and normal brazing in air where a flux has to be used is also inappropriate. However, brazing methods can be carried out in a controlled atmosphere or vacuum environment without the need for a flux and these will give air-tight joins suitable for ultrahigh vacuum applications. Basically the method involves holding the components together in a jig mounted in a suitable vessel with the brazing metal around the

join. The components are then heated up to the melting temperature of the brazing metal, which should be significantly lower than the melting point of the components, either by r.f. induction heating or by a furnace. Copper, silver, gold, palladium and nickel can all be used as elements in the brazing alloys which are designed for joining specific metals. They can be applied to the join as wire or foil or as powder or paste. Normally, for ultrahigh vacuum components, the brazing is carried out in a high vacuum furnace capable of attaining temperatures up to 1400°C. Stainless steel can be joined using nickel-based alloys. A feature of these fillers is that the material in the molten state alloys with the stainless steel effectively raising the remelt temperature. This enables a higher service temperature to be used, for example for degassing, and also makes it possible to perform further subsequent brazing operations. For a good brazed join the component must be accurately placed in position with low fitting tolerances so that the molten metal will be contained in the gap by capillary action. However, it is possible to use the brazing alloys with special gap fillers which enable larger clearances, up to 1.5 mm, between the components.

2.6.4 Ceramic seals

Ceramics can be sealed together using a glass soldering technique and special solders, both devitrifying and non-devitrifying, have been developed for this purpose. For the ceramics with a glassy phase there are no problems in 'wetting' the surface to get a good vacuum-tight seal but for the high purity alumina ceramic it is more difficult. Even if a devitrifying type of solder glass is used the technique does limit the service temperature of the ceramic and to some extent negates the value of using ceramics instead of glass.

An alternative method is to metallize the ceramic and to metal braze the components in a similar way to joining ceramic to metal described later. Such joints will withstand temperatures above 1000°C.

High purity alumina is difficult to metallize, particularly if a high-temperature seal is required. Klomp and Botden[47] showed, however, that a strong high-temperature seal could be obtained by using a ceramic solder of oxides from the system $Al_2O_3 . CaO . MgO . SiO_2$ with melting temperatures above 1200°C.

Some ceramics can be joined to glass directly by fusing the glass on to it. The expansion coefficients must be fairly closely matched which is difficult since ceramics do not show the increase in expansion coefficient at the softening point of the glass. Low expansion coefficient glass and ceramics are best matched particularly if the ceramic has a large amount of glassy phase. It is difficult to obtain a satisfactory glass seal to alumina. More usually a glass soldering technique is used.

The main use of ceramics in ultrahigh vacuum systems however is for the insulation of electrical feed throughs and therefore sealing of ceramic to metal is the prime interest. Once again the pioneering work in developing suitable techniques was carried out in the electronics industry in the 1940s, particularly for microwave tubes. As with glass, the most satisfactory seals are those between a metal and a ceramic with similar thermal expansion coefficients over a wide temperature range. Since the ceramic is much stronger than glass the match does not need to be so exact. In *Figure 2.10* the thermal

expansion characteristics were given for three or four ceramics and the metals to which they can be sealed were also shown. It will be seen that in the case of alumina, the nickel–iron thermal expansion differs substantially at the higher temperatures. However the strength of the alumina and resistance to thermal shock, coupled with the malleability of the nickel–iron, permits successful sealing in spite of the thermal expansion differences.

The normal method of sealing the ceramic to the metal is to first metallize the ceramic and then to braze it to the metal using the same materials as used in brazing two metal components together (see Section 2.6.3).

The key to a successful vacuum-tight brazed join of this type is the production of a good adhering metal layer on the ceramic part. This has been the subject of considerable study and several methods have been evolved. The two main methods are the sintered metal technique and the active metal technique. In the former method finely divided metal powder such as tungsten, molybdenum, iron or nickel in a suitable suspension of nitrocellulose is applied by brushing or other means to the ceramic as a band in the area to be joined. The metal coating is then sintered at a high temperature, 1300–1600°C depending on the ceramic, in a hydrogen atmosphere. A further coating of Ni or Cu is usually applied by electroplating and the ceramic is then brazed to the metal part. Mixing manganese powder with the other metals mentioned above improves the bonding and is generally used today. The main mixture used for alumina is molybdenum + manganese in the ratio of about 4:1 by weight. It is considered that the manganese is oxidized and forms a low melting phase with the SiO_2 of the ceramic which then penetrates the sintered molybdenum particles to form a strong bond. In general, therefore, the system only works when there is at least 1.5% SiO_2 glassy phase present in the ceramic. Some metallizing paints include SiO_2 in the powder and are claimed to be satisfactory for the high purity (>98%) alumina.

It had been found that active metals such as zirconium, tantalum and titanium when melted in contact with a ceramic form an intimate chemical bond. The second method depends on these findings. The metal powders themselves have a very high melting point, ~1700°C, and to lower the bonding temperature they are either applied as a lower melting point alloy or more usually applied as the metal hydrides. The normal procedure is to apply a slurry of titanium or zirconium hydride to the ceramic surface and heat to decomposition temperature, around 600°C, in vacuum. The ceramic with the metal layer is then brazed to the metal component in the normal way. It is during the brazing that the chemical reaction takes place to form the bond and at the same time the titanium alloys with the brazing metal. The reaction does not require the presence of SiO_2. Although commercially exploited, the strength of the seal is not considered to be as good as for the sintered metal technique.

For very high purity alumina, neither process is considered completely satisfactory and Klomp and Botden[47] advocate their oxide process with metal oxide powder such as molybdenum oxide added to the mixture (80% MO added to 20% of the oxide mix by weight). Other methods of providing a metallized ceramic have been advocated, for example Bronnes *et al.*[48] have produced a good metallized layer by a sputtering technique. For a more detailed review of the various techniques the reader is referred to the Chapter devoted to ceramic seals in the handbook by Kohl[49].

Whatever metallizing process is used, it is essential that the components to be joined are outgassed and clean with no mechanical faults such as cracks or fissures. Choice of the brazing eutectic is also important especially for the active metal coatings, since, for example, titanium and nickel can form a brittle intermetallic compound.

Ceramics can also be sealed to metal by thermal compression bonding. RCA[50] developed a ram seal whereby a silver-plated Iconel sleeve was forced over the tapered end of a ceramic tube to give a vacuum seal. Other workers[51] have experimented with pressure seals at elevated temperature in a vacuum or inert gas environment with some success. The method is limited, however, to butt or disc seals and is expensive.

Finally, ceramics can be sealed to metal with a solder glass, particularly the devitrifying types which have a high melting temperature.

2.6.5 Sealing other materials

For some applications special windows are required in the ultrahigh vacuum system. Examples of such windows are mica, sapphire, quartz, lithium fluoride, magnesium fluoride and germanium. In general, the windows must not be deformed, so that sealing at a temperature well below the softening point of the window material must be employed. The main methods of sealing are, therefore, glass soldering or metal brazing techniques but other methods such as thermo-compression bonding with a soft metal interface may be appropriate.

Mica windows have in the past been sealed with silver chloride. Silver chloride has some unusual properties which allow vacuum sealing when the more normal techniques cannot be employed. It has a melting point of 457°C and when molten it will wet mica and precious metals. It reacts however with the common metals, being decomposed by a replacement reaction. When cooled it is a wax-like compound which behaves as an elastic material at low stress and becomes ductile at higher stress. It is thus able to accommodate variations in the expansion coefficients of the two components to be joined. Chemically it is not particularly stable. Although it is insoluble in water and most acids, it is soluble in NH_4OH and $Na_2S_2O_3$. It is affected by UV radiation and is easily scratched. A more satisfactory seal can be obtained using low-temperature soft solder glass, in particular those that melt below the onset of calcination which is around 650°C for natural mica and 900°C for synthetic mica. Mica has a thermal expansion coefficient parallel to the cleavage plane of around $90 \times 10^{-7}°C^{-1}$ and bonds readily to soft glass. Synthetic mica, because of its higher calcinating temperature, can also be metallized and brazed. It is essential that the sealing material, whether it be silver chloride, glass frit or brazing metal, covers the edges of the window to prevent the mica splitting in use.

Sapphire and quartz, both having high-temperature melting points and resistance to thermal shock, can be joined to metal using a solder glass or a metal braze in a similar way to ceramic. To allow for differences in the expansion coefficients of the window and the metal holder it is advisable to thin the metal at the join. An example of a window mounted in a flange is illustrated in *Figure 2.14*. Both sapphire and quartz can be sealed directly to glass using the normal glass blowing techniques of fusing the glass on to the

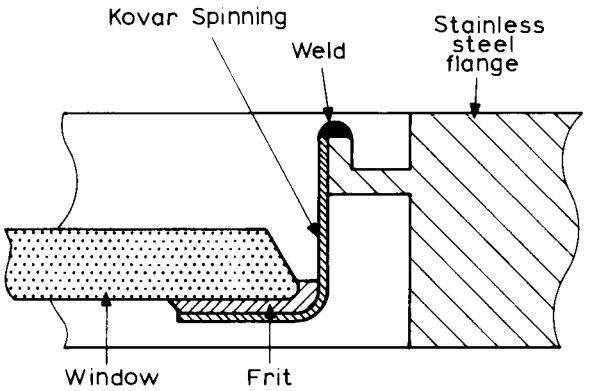


Figure 2.14 A sapphire window mounted in a flange

window. Sapphire with a thermal expansion coefficient of 58 to 90 $\times 10^{-7}°C^{-1}$ is reasonably matched to the soda glasses. Corning has a glass No. 7530 with an expansion coefficient of 71 $\times 10^{-7}°C^{-1}$ which is an especially good match. Quartz requires a graded seal technique to connect it to a metal sealing glass.

For the lower melting point materials, glass soldering techniques can usually be employed. There are now a range of glass solders covering a wide range of thermal expansion coefficients. Magnesium fluoride can be sealed in this way[52]. Mulder[53] describes a bakable seal for a lithium fluoride window using lead fluoride to copper and gold.

Where thermal expansion coefficients cannot easily be matched, thermo-compression seals using an intermediate soft metal such as indium or lead can often be used. Germanium windows are often attached to metal flanges in this manner. The low melting points of the metals restrict the baking temperature, but in general they are more satisfactory from the vacuum point of view than epoxy seals. To obtain good bonds it is sometimes necessary to plate or otherwise coat the sealing surface of the window with a suitable metal. For example for germanium an evaporated layer of nichrome followed by gold makes an excellent bond to indium.

2.7 References

1. DOUGLAS, R. W., *J. Sci. Instrum.*, **22**, 81, (1945)
2. DUSHMAN, S., *Scientific Foundations of Vacuum Technique*, 2nd Edit., Wiley New York, (1962)
3. ALPERT, D. and BURITZ, R. S., *J. Appl. Phys.*, **25**, 202, (1954)
4. NORTON, F. J., *J. Am. Ceram. Soc.*, **36**, 90, (1953)
5. ALTEMOSE, V. O., *J. Appl. Phys.*, **32**, 1309, (1961)
6. ALPERT, D., *Advances in Vac. Sci. Technol. Vol. 1*, p. 31. Proc. 1st Internat. Congr. on Vac. Tech. 1958, Pergamon Press, (1960)
7. TODD, B. J., *J. Appl. Phys.*, **26**, 1238, (1955)
8. HONIG, R. E., *RCA Review*, **23**, 567, (1962) and *ibid*, **30**, 285, (1969)
9. BLEARS, J., GREER, E. J. and NIGHTINGALE, J., *Advances in Vac. Sci. Technol. Vol. 2*, p. 473, Proc. 1st Internat. Congr. on Vac. Tech. 1958, Pergamon Press, (1960)
10. DAYTON, B. B., *Trans 8th Nat. Vac. Symp. and 2nd Internat. Congr. on Vac. Technol., 1961 Vol. 1*, p. 42 Pergamon Press, (1962)
11. REITER, F. and CAMPOSILVAN, J., *Vacuum*, **32**, 227, (1982)

12. ESPE, W., *Materials of High Vacuum Technology Vol. 1* (English Trans), Pergamon Press, (1966)
13. NORTON, F. J. and MARSHALL, A. L., *Trans. Am. Inst. Min. Metall. Engr.*, **156**, 351, (1944)
14. FLECKEN, F. A. and NÖLLER, H. G., *Trans 8th Nat. Vac. Sci. Technol. and 2nd Internat. Congr. on Vac. Sci. Technol., 1961, Vol. 1*, p. 58, Pergamon Press, (1962)
15. VARADI, P. F., *Trans. 8th Nat. Vac. Symp. and Proc. 2nd Internat. Congr. on Vac. Sci. Technol., 1961 Vol. 1*, p. 73, Pergamon Press, (1962)
16. GOVIER, R. P. and McCRACKEN, G. M., *J. Vac. Sci. Technol.*, **7**, 552, (1970)
17. JONES, A. W., JONES, E. and WILLIAMS, E. M., *Vacuum*, **23**, 227, (1973)
18. HOLLAND, L., *Vacuum*, **26**, 97, (1976)
19. NORTON, F. J., *Trans. 8th Nat. Vac. Symp. and Proc. 2nd Internat. Congr. on Vac. Sci. Technol., 1961, Vol. 1*, p. 8, Pergamon Press, (1962)
20. PERKINS, W. G., *J. Vac. Sci. Technol.*, **10**, 543, (1973)
21. ESPE, W., *Materials of High Vacuum Technology, Vol. 2* (English Edit.), Pergamon Press, (1968)
22. GROSSMAN, D. G., *Vacuum*, **28**, 55, (1978)
23. MILLER, C. F. and SHEPARD, R. W., *Vacuum*, **11**, 58, (1961)
24. ALTEMOSE, V. O. and KACYON, A. R., *J. Vac. Sci. Technol.*, **16**, 951, (1979)
25. HAYES, D., BUDWORTH, D. W. and ROBERTS, J. P., *Trans. Br. Ceram. Soc.*, **62**, 507, (1963)
26. BUDWORTH, D. W., *Trans. Br. Ceram. Soc.*, **62**, 975, (1963)
27. GIBBONS, W. F., *Proc. 4th Internat. Congr. on Vac. Sci. Technol. 1968 Inst. Phys. Conf. Series 5 Part 1*, p. 255, (1968)
28. NORTON, F. J., *J. Appl. Phys.*, **28**, 34, (1957)
29. GELLER, R., *Le Vide*, **13**, 71, (1958)
30. DIELS, K. and JAECKEL, R., *Leybold Vacuum Taschenbuch*, Ch. 13, Springer Berlin (1958), (English translation, Pergamon Press, (1966))
31. SANTELER, D. J., *Trans. 5th Nat. Symp. Vac. Technol.*, p. 1, (1958)
32. OTHMER, D. F. and FROHLICH, G. J., *Ind. Eng. Chem.*, **47**, 1034, (1955)
33. BARTON, R. S. and GOVIER, R. P., *J. Vac. Sci. Technol.*, **2**, 113, (1965)
34. DAYTON, B. B., *Trans. 6th Nat. Symp. Vac. Technol.*, p. 101, (1959)
35. HAIT, P. W., *Vacuum*, **17**, 547, (1967)
36. DE CHERNATONY, L., *Vacuum*, **27**, 605, (1977)
37. MARKLEY, F., ROMAN, R. and VOSECEK, R., *Trans. 8th Nat. Sym. Vac. Technol. and Proc. 2nd Internat. Congr. on Vac. Sci. Technol., 1961 Vol. 1*, p. 78, Pergamon Press, (1962)
38. KENDALL, B. R. F., *J. Vac. Sci. Technol.*, **20**, 248, (1982)
39. BAILEY, J. R., *Handbook of Vacuum Physics Vol. 3, Part 4*, Pergamon Press, (1964)
40. BRUBAKER, D. W. and KAMMERMEYER, K., *Ind. Eng. Chem.*, **44**, 1465, (1952); **45**, 1148, (1953); **46**, 733, (1954)
41. GEORGE, D. E., Private communication quoted by Perkins[20]
42. PEACOCK, R. N., *J. Vac. Sci. Technol.*, **17**, 330, (1980)
43. LEWIN, G. and MARK, R., *Trans. 5th Nat. Vac. Symp.*, p. 44, (1958)
44. BENBENEK, J. E. and HONIG, R. E., *Rev. Sci. Instrum.*, **31**, 460, (1960)
45. WALLIS, G., DORSEY, J. J. and POMERANTZ, D. I., *Proc. 14th Symp. on Art of Glassblowing*, The American Scientific Glassblowers Society, p. 47, (1969)
46. KRONBERGER, H., *Proc. I.M.E.*, **172**, 113, (1958)
47. KLOMP, J. T. and BOTDEN, Th. P. J., *Am. Ceram. Soc. Bull.*, **49**, 204, (1970)
48. BRONNES, R. L., HUGHES, R. C. and SWEET, R. C., *Philips Tech. Rev.*, **35**, 209, (1975)
49. KOHL, W. H., *Handbook of Material and Techniques for Vacuum Devices*, Reinhold Publishing Corp., New York, (1967)
50. MARTIN, I. E. and TUNIS, A. C., *RCA Eng.*, **3**, 9, (1957)
51. McNALLY, J. O., *Proc. 1958 Electronics Corp. Conf.*, Engineering Publishers New York, p. 168, (1958)
52. FREEMAN, G. H. C. and MOORE, P. J., *J. Phys. E. Sci. Instrum.*, **11**, 980, (1978)
53. MULDER, B. J., *J. Phys. E. Sci. Instrum.*, **10**, 591, (1977)

Chapter 3

Pumps

3.1 Background

In Chapter 1 we saw that the ultimate pressure attained in a vacuum system depends on the influx of gas on the one hand and the pumping of the gas on the other. Whereas the gas influx can vary over several orders of magnitude and can be reduced greatly by the selection of the materials used in the vacuum system and its processing, the pumping rate is more restricted. There is, for example, the theoretical limit determined by the rate at which the gas molecules are incident on the pump orifice (see Section 1.4.2). The ultimate pressure cannot, therefore, be lowered indefinitely by increasing the size of the pump beyond the orifice and to decrease the pressure by an order, without changing the influx of gas, would require the conductance of the orifice to be increased by an order as well as the pump speed.

In practice, the speed of a pump is less than the theoretical limit and will depend on the pressure of the gas being pumped. Indeed, most pumps only work over a limited pressure range and outside this range the pumping speed falls to zero. Of first importance, therefore, in considering the evacuation of the gas from a system is the selection of a suitable pump, which, since there is no pump covering the whole pressure range from atmosphere to ultrahigh vacuum, implies the selection of the best combination of pumps. The size of the pump is then dictated by the volume of the system and the allowable orifice.

Until the 1950s the only pumps capable of reaching pressures below 10^{-5} Pa were diffusion pumps, backed by mechanical rotary pumps to give a backing pressure of about 1–10 Pa. Since there was little need for vacuum systems of lower pressures than 10^{-4} Pa, very little attention was paid to the efficiency of the pumping system or the accuracy of the pressure measurements. The need for better vacuum for surface studies, space simulation and particle accelerators, for example, saw the upsurge of research and development work on the attainment and measurement of vacuum in the 1950s and 1960s. This resulted not only in the improvement of the diffusion pump/rotary pump system, extending its range down to ultrahigh vacuum pressure but also in the introduction of several alternative pumps and pumping techniques. These techniques were basically not new and indeed similar systems were employed before the advent of the diffusion pump in 1915. However, a re-

examination of these early techniques in the light of the better appreciation of the principles involved, resulted in the development of sophisticated pumps which can evacuate efficiently large systems to ultrahigh vacuum. They include ion pumps, sublimation pumps, sorption pumps and cryogenic pumps. Such pumps depend on trapping the gas either physically or chemically within the body of the pump and not only is a combination of pumps required to cover the full range down from atmospheric pressure but, because the mechanism can be selective, a combination may also be required to cover the gas species.

Another modern pump which is becoming increasingly popular is the turbomolecular pump. Based on the molecular drag pump of Gaede[1] in 1913, it has been developed to give high pumping speeds down to ultrahigh vacuum pressures of 10^{-8} Pa.

This chapter describes these various pumps and discusses their performance and application to ultrahigh vacuum systems.

3.2 Diffusion pumps

Diffusion pumps are adequately described in most vacuum technology handbooks (see for example Pirani and Yarwood[2]) and no special design from the fundamental point of view is necessary for ultrahigh vacuum application. However, before considering the requirements for achieving low pressures it is useful to outline the mechanism of a diffusion pump.

The operation of the pump is based on the diffusion of the gas in a stream of vapour. The vapour stream is produced in the pump by boiling a suitable fluid (oil or mercury). The vapour passes up a chimney-like structure and through an annular jet which directs the stream back down the outside of the chimney, see *Figure 3.1*. The density of the vapour at the jet is much greater than that of the pumped gas and the gas molecules are entrained in the vapour stream. The vapour is then condensed on the water-cooled walls of the pump and the gas, which now has a greater density than at the pump orifice, passes on to a second jet stage where the process is repeated. One or more other stages may follow, the gas being ultimately pumped away by the rotary backing pump. Modern pumps usually have a side jet or ejector stage into the backing pump to increase the backing pressure tolerance (*Figure 3.1*). The upper pressure limit of the pumping mechanism is set by the requirement that the gas pressure must be lower than the vapour pressure of the heated pump fluid at the jets, but there is no fundamental reason why it should not continue to operate down to very low pressures.

Reappraisal of the diffusion pump in the 1950s suggested that other factors were limiting the ultimate pressure and it was the study of these effects which now enables lower pressures to be achieved and makes the diffusion/rotary pump combination the main contender for pumping ultrahigh vacuum systems today.

The factors that can affect the ultimate pressure of a diffusion pump are listed as follows:

(1) Back diffusion of the pumped gas against the vapour stream.
(2) Saturation vapour pressure of the pump fluid or of decomposition products.

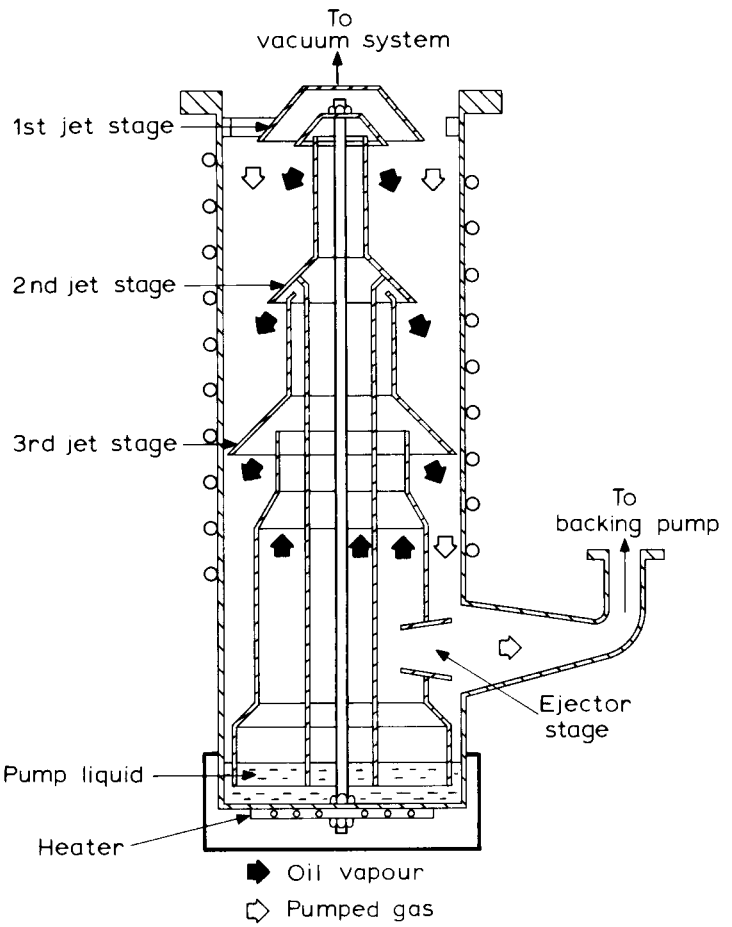


Figure 3.1 Sectional diagram of a three-stage diffusion pump

(3) Evolution of gas from the pump components.
(4) Gas becoming dissolved in the pump fluid and released when heated.

Back diffusion results in the setting up of a constant ratio between the pressure at the outlet and that at the inlet. Analogous to the mechanical pump, this is termed the compression ratio. The compression ratio in the case of a diffusion pump depends on the molecular weight of the gas being pumped. It is lowest for the light gases, such as helium and hydrogen, and increases markedly for the heavier gases, such as nitrogen. However, it is not difficult to design diffusion pumps with a compression ratio for hydrogen of 10^4 and that for nitrogen of 10^{10} and modern designs have been developed[3] to give such ratios. With these compression ratios, ultimate pressures well into the ultrahigh vacuum region should be expected with modest backing pump pressures, provided large quantities of hydrogen or helium are not being pumped. Thus, although attention must be paid to the compression ratio, it is clear that the demands placed on the design on on this score are not

particularly severe for ultrahigh vacuum application and most modern pumps will meet the requirements.

The saturation vapour pressure of the pump fluid is a more serious problem. Although there are pump fluids which have vapour pressures of the order of 10^{-8} Pa at room temperature, such as 705 silicone oil and polyphenyl ether, there is always the possibility of decomposition of the fluid into higher vapour pressure components or that it does not condense until it reaches the vacuum chamber.

Two effects can be distinguished when considering this problem. (1) back-streaming, whereby some of the vapour molecules travel in the wrong direction at the jets and are accelerated through the pump orifice and (2) back migration, which is the re-evaporation of fluid which has condensed around the orifice region. Back-streaming can be a very large effect, especially in oil pumps, and results in a vapour pressure in the vacuum chamber greatly in excess of the saturation vapour pressure at the wall temperature. It can also cause a serious loss of fluid from the pump. The effect can be reduced by several orders with suitably designed baffles placed in the throat of the pump, on which the pump fluid can be condensed and be returned to the pump. Power and Crawley[4] measured a 3000:1 reduction in the amount of oil back-streaming in one of their pumps when a copper disc was mounted above the top jet. Infact, since a large proportion of the back-streaming occurs at the top jet, a 'cold-cap' just above the jet is incorporated in most modern commercial pumps. The most efficient baffle however is that known as the chevron baffle, illustrated in *Figure 3.2*. It consists of a series of metal strips bent in a 'V' shape along their length and mounted parallel to one another and close enough to

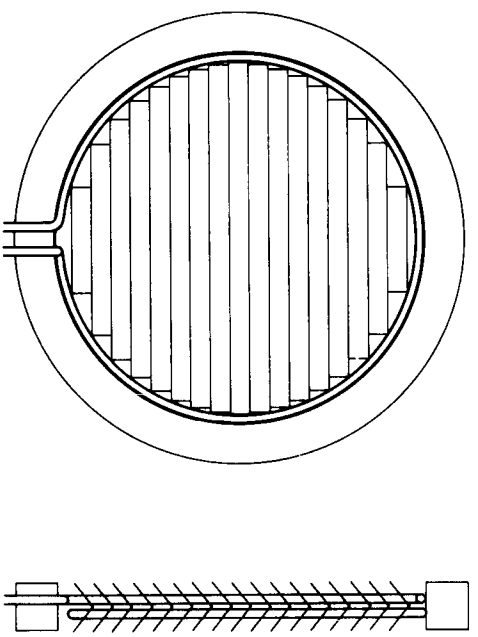


Figure 3.2 Typical construction of a water-cooled chevron baffle

ensure that there is no straight optical path through the baffle. A molecule passing through the baffle must then hit the surface at least once whatever its initial direction. Unfortunately, the more restrictive the baffle is to back-streaming the more it will reduce the effective pump speed and, according to Riddiford[5], the chevron baffle reduces the pump speed by about 50%.

The baffle in itself does not prevent migration but in general it will be at a lower temperature than, for example, the jet assembly, so that back migration will also be reduced. A further improvement can be obtained by cooling the baffle with water or a coolant from a refrigerator.

A cooled chevron baffle combined with a low vapour pressure oil could be adequate for many ultrahigh vacuum applications; Crawley *et al.*[6] recorded an ultimate pressure of 6.7×10^{-7} Pa with such a system using 705 silicone oil. For higher vapour pressure oils and certainly mercury, a cooled baffle is not sufficient and a liquid nitrogen cooled trap is required. At liquid nitrogen temperatures the vapour pressures are well below ultrahigh vacuum requirements but the trap must be designed to ensure that the chances of molecules passing through it without hitting the refrigerated surface are negligible. Several designs of trap have been proposed in the literature and *Figure 3.3(a)* illustrates a fairly typical design for metal systems, which is mounted directly on the top of the pump. *Figure 3.3(b)* illustrates a trap for a glass ultrahigh vacuum system designed by Venema and Bandringa[7] which gives a high conductance, 1.5×10^{-2} $m^3 s^{-1}$, yet a high condensation efficiency. Since the

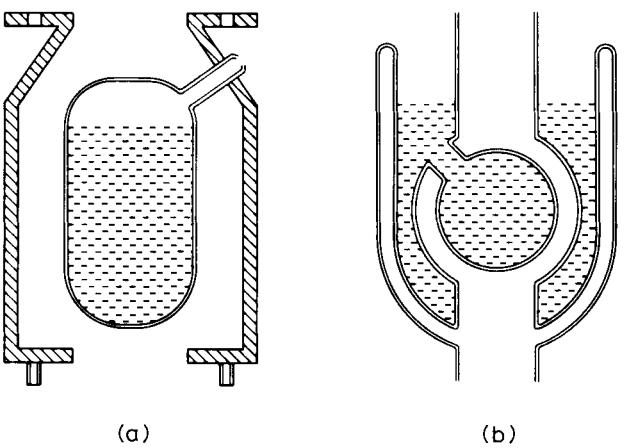


Figure 3.3 (a) Metal cold trap; (b) Glass cold trap according to Venema and Bandringa[7]

vapour condensed in the trap is lost to the pump a cooled baffle is still required to reduce back-streaming into the trap and return the bulk of the oil or mercury to the pump.

The combination of a water cooled baffle and a liquid nitrogen trap is therefore usually necessary for ultrahigh vacuum systems, but it should be realized that such a combination will restrict the pump speed probably to less than 40% of its potential. They also present large surface areas from which other gases may be evolved.

This brings us to the third factor limiting the ultimate pressure, namely gas evolved from the pump components. Any gas evolved within the pump, especially in the region of the pump orifice, will tend to counteract the pumping action. The problem is the same as for other components in the vacuum system and the considerations which apply to materials for ultrahigh vacuum use also apply to the pump components. The walls must be impermeable to atmospheric gases and the components should have a low outgassing rate. Since the degassing rate will be considerably reduced by baking, the pump, the trap and the baffle should be designed to allow bakeout to at least 250°C especially in the region above the jets.

Commercial diffusion pumps today are constructed with stainless steel bodies and nickel plated copper or steel jets, which are satisfactory for ultrahigh vacuum. However, connection between the pump and trap and to the system is normally via an elastomer 'O'-ring. Even if a bakable gasket such as Viton 'A' is employed for these seals, the pump is rarely designed to allow a baking cycle. On the other hand most manufacturers now offer modified designs for ultrahigh vacuum systems, where a bakable flange with a metal gasket replaces the 'O'-ring seal. With this modification the top area of the pump can be baked to 250°C and the outgassing rate reduced to an acceptable level. A modern design from Edwards High Vacuum, the Diffstak, illustrated in *Figure 3.4*, combines the pump with a watercooled baffle in one unit. The design reduces the number of seals and the amount of metal that requires degassing. Because the baffle sits in the wider section of the pump, the pumping speed is less restricted than in other designs. The design is available without the butterfly valve and with a metal gasket seal.

The remaining source of gas within the pump, i.e. the gas dissolved during condensation in the pump fluid and subsequently released when it becomes reheated, is more difficult to eliminate. A cooled ejector stage is purported to give efficient degassing of the condensing liquid, but little evidence is available to support this idea. It may well be, therefore, that this process determines the ultimate pressure attainable with diffusion pumps and virtually nothing can be done to reduce it. Gases in general are more readily dissolved in oil than in mercury so this effect should be minimal with mercury diffusion pumps. On the other hand, the ultimate pressures with mercury or oil pumps are very similar, provided that a liquid nitrogen trap is used in the case of the mercury pump, necessary since the vapour pressure of mercury at room temperature is around 10^{-1} Pa.

The choice of pump fluid is an important consideration and depends largely on the application. Because of the high vapour pressure at room temperature, the relative sensitivity to contamination and the health hazard of mercury, it is rarely used today. It is then a choice of the type of oil. There are several types available for diffusion pumps and most pumps will work satisfactorily with any one of them. Oils are complex compounds and even the most stable can fractionate or decompose slightly when heated or subjected to electron bombardment. The early mineral oils, the best known being Apiezon oils, were prone to oxidation and accidental exposure to air could leave tar-like deposits in the pump. The introduction of silicone oils in the 1950s overcame this problem. Such fluids are thermally stable, do not oxidize and in fact have a high resistance to chemical attack. Dow Corning produce a range for different ultimate pressures, the one with the lowest vapour pressure of around 10^{-7} Pa

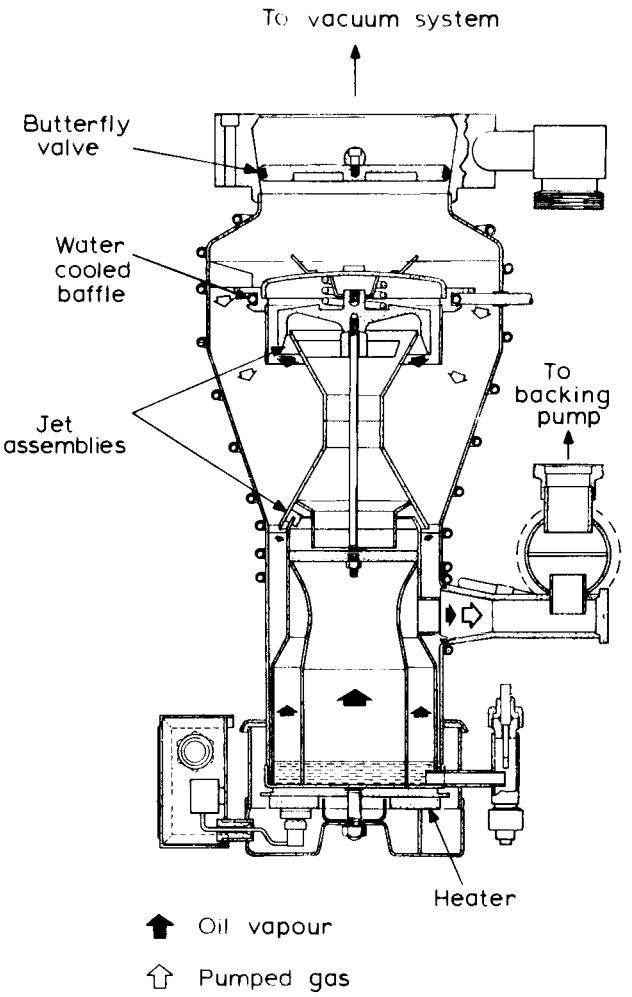


Figure 3.4 Section through the Diffstak oil diffusion pump. (Courtesy Edwards High Vacuum)

being DC 705. The main disadvantage of silicone oils is that they can be polymerized under electron bombardment to form insulating films on electrodes.

In recent years a number of man-made fluids such as esters, ethers and naphthalenes have been developed for diffusion pump use. The best known are polyphenyl ether and perfluoro polyether. Polyphenyl ether, as a blend of isomers, is marketed under various trade names, the best known being Santovac 5 from Monsanto. It has a vapour pressure at 20°C lower than 10^{-7} Pa and when polymerized under electron bombardment forms a conducting film. It is, however, very expensive. Perfluoro polyether has a higher vapour pressure $\sim 10^{-6}$ Pa but does not polymerize under electron bombardment. Both fluids are thermally and chemically stable, although the perfluoro polyether will decompose above 300°C. Edwards market a naph-

TABLE 3.1. Comparative properties of diffusion pump fluids

Fluid	*Apiezon C*	*Silicones*		*Polyphenyl ether*	*Perfluoro polyether*	*Edwards L9*
		702	*705*			
Vapour pressure at 20°C Pa	4×10^{-7}	6.5×10^{-5}	2.6×10^{-8}	2.6×10^{-8}	2.7×10^{-6}	7.8×10^{-8}
Ultimate pressures obtained with Diffstak Pa	1.3×10^{-7}	6.5×10^{-4}	1.3×10^{-7}	$<10^{-7}$	3.0×10^{-6}	5×10^{-7}
Effect of energetic particles	Conducting polymer	Insulating polymer		Conducting polymer	No polymer except with H_2 ions	Conducting polymer
Thermal stability	Poor	Very good		Excellent	Decomposes above 300°C to gas	Good
Oxidation resistance	Poor	Excellent		Very good	Excellent	Very good
Chemical resistance	Poor	Very good		Relatively low	Excellent with a few exceptions	Good
Cost	Low	Low	Moderate	High	High	Low

thalene-based oil under the code L9 which has similar properties to polyphenyl ether. A summary of the main properties of these oils is given in *Table 3.1* taken from a paper by Laurenson[8] to which the reader is referred for more detailed information.

A factor which is not always appreciated when using a diffusion pump–rotary pump combination is the possibility of oil from the rotary pump entering the diffusion pump and being heated; the problem is discussed by Holland[9]. The rotary pump oil is chosen for its lubricating properties as well as its vacuum behaviour. It is quite unsuitable for diffusion pumps in terms of vapour pressure and chemical stability. Baffle valves and water-cooled traps available for connecting between the rotary pump and diffusion pump are unsatisfactory for ultrahigh vacuum systems. Even liquid nitrogen traps, because they are less effective under the viscous flow conditions met with at the backing pressure, are not completely safe. The best trap is that containing a ceramic based absorbent such as zeolite.

Zeolite readily absorbs vapours and gases because of its very porous surface (see Section 3.4) especially hydrocarbons and water vapour. The adsorption increases with decreasing temperature and foreline traps of this type are designed for liquid nitrogen cooling. The main problems with this type of trap are that the surfaces become saturated after a time, reducing its absorbing power and also, if a large amount of water vapour is pumped, the trap can be blocked by ice. It can be regenerated, however, by baking and can be bypassed during the initial degassing period. Baker and Laurenson[10] suggest that activated alumina is better than zeolite in these respects. In conclusion, we may say that, provided the pump fluid vapour is adequately trapped and the top of

the pump with baffle and trap can be baked to at least 250°C, modern diffusion pumps can be employed to achieve ultrahigh vacuum conditions down to pressures of the order of 10^{-8} Pa. However there is always a possibility of slight oil vapour contamination which may be unacceptable for some applications.

3.3 Turbomolecular pumps

The turbomolecular pump is a development of the early molecular drag pump first introduced by Gaede[1] in 1913. In its simplest design, the molecular drag pump consisted of a metal cylinder which could be rotated at high velocities in a vacuum sealed casing, illustrated in *Figure 3.5(a)*. Mechanically, it closely resembled the rotary pump, but differed in having no mechanical separation between the fine and the rough vacuum ports. The gas molecules entering the pump with thermal velocities have a high probability of hitting the moving cylinder surface, where they will dwell for a finite time. When they are released from the surface they will have an extra velocity in the direction of the cylinder rotation. Thus they are dragged through the pump by the rotating cylinder receiving impulses at each collision. Essential criteria for the functioning of the pump are, that the impulse received should be significant in relation to the thermal velocity of the molecules and that their mean free path should be large compared with the pumping area to keep intermolecular collisions to a minimum. Gaede showed that under these conditions of molecular flow, the ratio of the pressure at the output to that at the input could be expressed by the following equation

$$\frac{P_{\text{out}}}{P_{\text{in}}} = \exp(Av) \tag{3.1}$$

where v is the angular velocity of the cylinder which needs to be of the order of 10^4 rev/min, and A is a constant determined by the geometry of the pumping area and the nature of the gas. To obtain a large value of A, the surface area of the cylinder exposed to the gas should be large, but the gap between the cylinder and casing small. Also the backing pressure must be less than 100 Pa to ensure molecular flow conditions.

In a practical design the surface area was increased by providing the rotating cylinder with vanes which passed close to interdigital vanes on the casing, *Figure 3.5(b)*. The rotating cylinder was about 5 cm diameter and the gap between cylinder and casing 0.1 mm. The sections were connected up in series to give a compression ratio of 10^5 for nitrogen. The pump speed however, was relatively low, of the order of 10^{-3} m^3s^{-1}. Various other designs have been suggested over the years, in particular the design of Holweck[11] in 1923 in which the cylinder was smooth but the casing was provided with a helical groove of varying depth so that the gas was pumped around and along the axis of the pump. A similar idea was exploited in a design devised by Siegbahn and described by von Friesen[12] in which a simple rotating disc is spun close to a surface in which are cut several spiral grooves along which the gas is pumped. However, because of the requirement for high rotational speeds, with close proximity between rotating and stationary components and

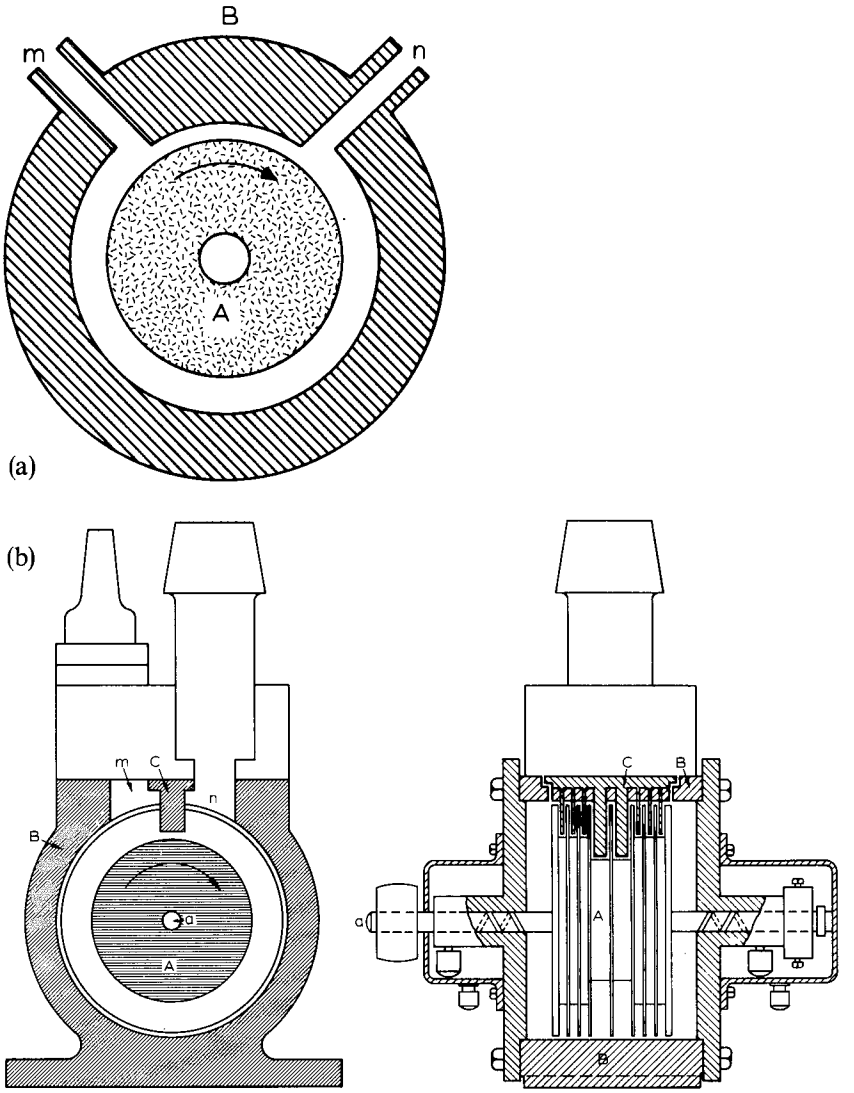


Figure 3.5 (a) Principle of operation of a Gaede molecular pump;
(b) Construction of practical molecular drag pump; a, axis; A, rotor; B, case;
C, stator; n, inlet; m, outlet

because of the rather poor pumping speed performance, none of these designs was commercially exploited.

The situation was changed with the development of the turbomolecular pump first described by Becker[13] in 1958. In this design the rotor uses a turbine form of blading with the interdigital vanes similarly bladed, as illustrated in *Figure 3.6(a)*. The important design feature is that the pump functions with a gap of millimetres between the rotor and stator vanes, allowing larger mechanical tolerances. The pumping action is also rather different. Although the gas is dragged round in the direction of the blades, it is

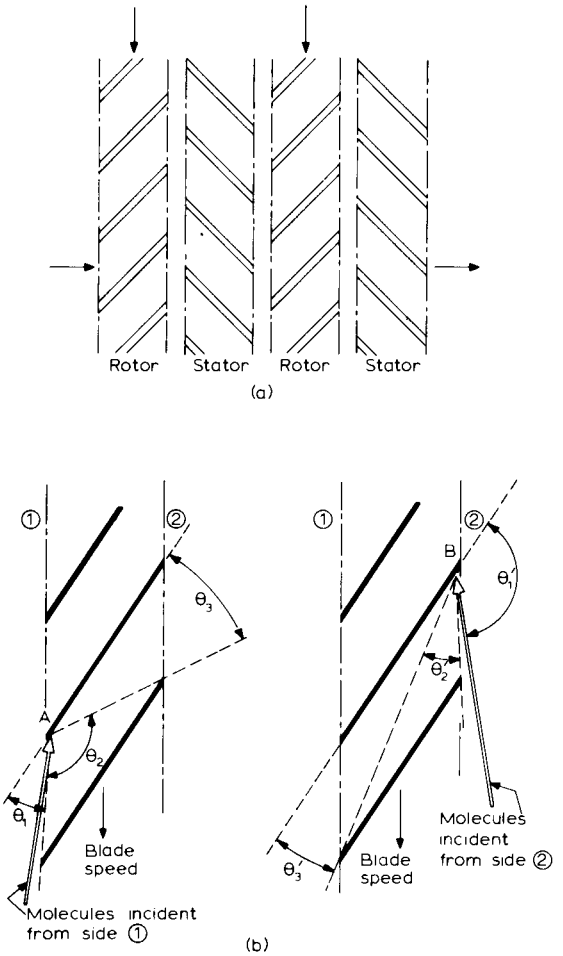


Figure 3.6 (a) Blade arrangement in the turbomolecular pump. (b) Diagram illustrating the pumping action of a turbomolecular pump. Blade speed is assumed large compared to molecular speed

actually pumped across the blades. The mechanism is illustrated in *Figure 3.6(b)*.

If we consider molecules of gas *en masse* drifting towards the rotating blades, then on average their relative velocities will be at a fairly sharp angle to the direction of rotation and they will impinge on the edges of the blades as illustrated. Assuming diffuse reflection, then on side 1 molecules re-emitted within the angle θ_1 will return to side 1, whilst those re-emitted within angle θ_3 will escape to side 2. Those re-emitted within angle θ_2 may ultimately escape either to side 1 or 2. A similar process will occur with molecules incident on side 2, the relevant angles being θ_1', θ_2', and θ_3' respectively. The probability that a molecule will transfer from one side of the blade to the other will depend on the ratio of the angles and clearly, from the diagram, the probability of a

molecule going from side 1 to 2 is much higher than for a molecule going from 2 to 1. The mechanism has been treated theoretically by Kruger and Shapiro[14] using Monte Carlo calculations. They showed that depending on blade speed and angle, the probability of molecules going from side 1 to 2 was some 10 to 40 times higher than for molecules going in the opposite direction.

The net flow of gas across the blades will not only depend on these probabilities, however, but also on the pressure ratio across the blades. Generally it was shown that the design for maximum pump speed gave the lowest compression ratio and vice versa, so that a compromise between compression ratio and pumping speed has to be chosen. The velocity the impinging molecules receive in the direction of the blade rotation means that they will also hit the stator blade at a glancing angle. Since the stator blades are angled in the opposite direction a similar favourable flow through the blades is obtained. In a practical design there will be a number of sets of vanes, each pair of rotor and stator sections representing a pumping stage. It is also fairly common in modern pumps to have some of the pumping stages designed for maximum pump speed whilst those nearest the outlet are designed to give a higher compression ratio. In the design described by Becker and manufactured by Arthur Pfeiffer GmbH, there were two sets of vanes which pumped the gas away from a central port; the design is illustrated schematically in *Figure 3.7*. The rotor speed depended on the size of the pump but was typically around 10 000 rev/min. This put a considerable strain on the bearings which had to be efficiently lubricated by a flowing oil system and water cooled. Pump speeds comparable with similar sized diffusion pumps were obtained, the Pfeiffer

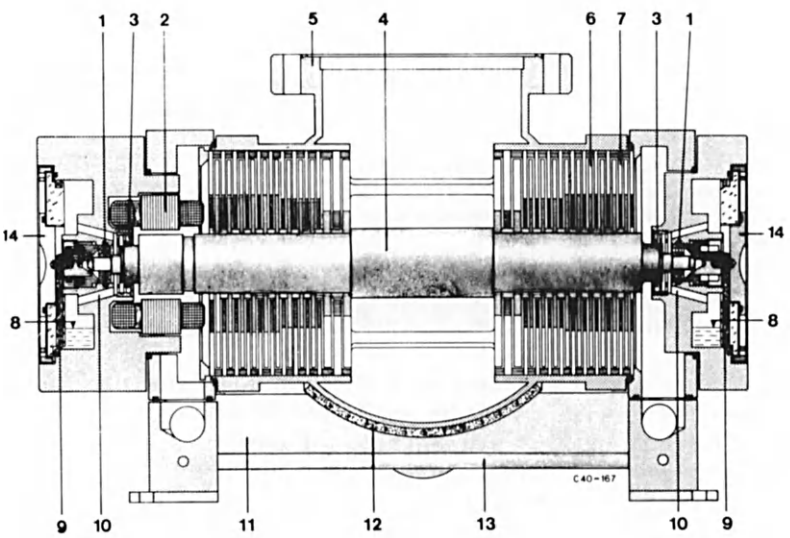


Figure 3.7 Section diagram of a turbomolecular pump based on the design of Becker[13]. 1, Bearing; 2, motor; 3, labyrinth chamber; 4, rotor; 5, UHV connection; 6, rotor vane; 7, stator vane; 8, oil reservoir; 9, oil supply for bearing; 10, oil return; 11, backing pressure channel; 12, heater; 13, water cooling; 14, seal. (Courtesy Balzers High Vacuum Ltd)

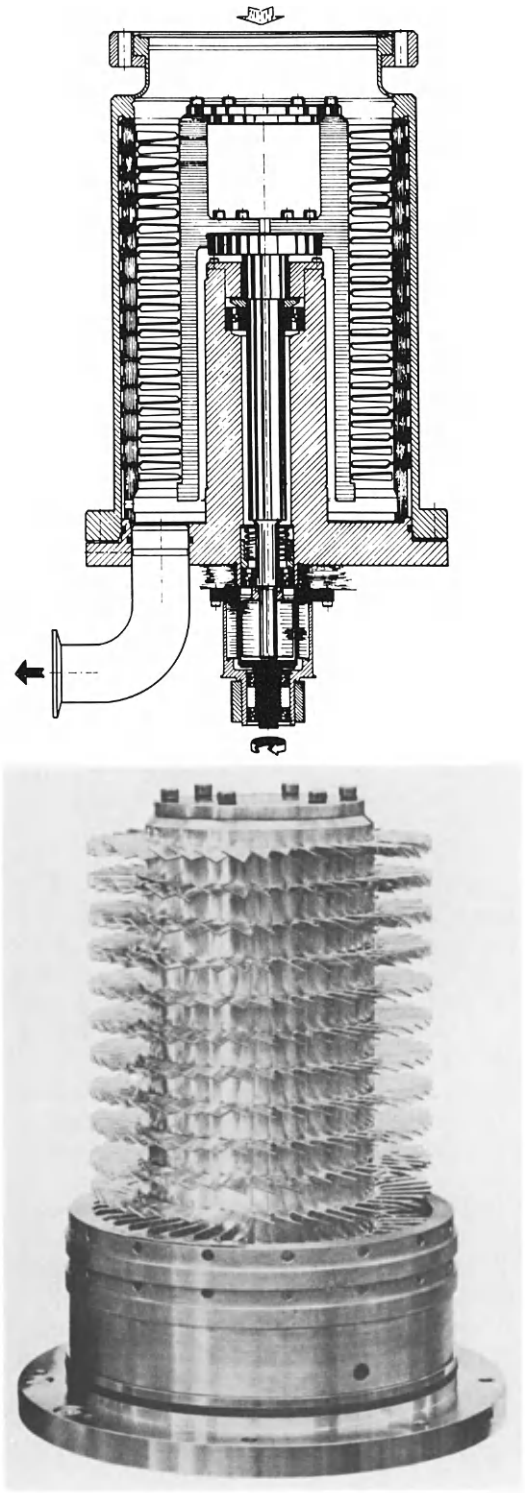


Figure 3.8 Sectional diagram and photogra
the motor of the turbomolecular pump acc
to Mirgel[16]

range covered pump speeds from $250\,m^3h^{-1}$ to $15\,000\,m^3h^{-1}$. The compression ratio was typically 10^9 for nitrogen and 10^3 for hydrogen.

Since the turbomolecular pump is backed by a rotary pump or sorption pump to ensure molecular flow conditions, residual pressures in the ultrahigh vacuum region, down to 10^{-8} Pa, can be achieved. The residual gases will be mainly the lighter gases with a predominance of hydrogen. Because of the high compression ratio for heavy molecules, oil vapour from the bearings does not back-stream and the pump is claimed to give a hydrocarbon free vacuum without the need for baffles or liquid nitrogen cooled traps. However, it must be pointed out that when the pump is not rotating, oil from the bearings can migrate into the vacuum system so that precautions must be taken against such eventuality.

The latest pumps have benefited from the application of the theory of Kruger and Shapiro and with redesigned blades, the pumps are more efficient than the original design of Becker. Indeed, by increasing the rotational speed to 42 000 rev/min, made possible by using transistor controlled d.c. motors, and sacrificing slightly on compression ratio, Osterstrom and Shapiro[15] have produced a design giving ten times the speed of a similar sized pump of earlier design, with only half as many pumping stages.

An alternative design has been described by Mirgel[16] in which the vanes are rotated about a vertical axis, with a unidirectional flow of gas, giving a more compact design than the horizontal pump. The design is shown diagrammatically in *Figure 3.8* and is marketed by Leybold–Heraeus. One of the advantages claimed is that there is no pressure differential at the bearing to drive lubricating fluid or vapours from the drive mechanism area to the vacuum chamber. The pump is approximately 20 cm diameter and 46 cm high and with a rotating speed of 24 000 rev/min gives a pump speed of $1332\,m^3h^{-1}$.

Leybold–Heraeus later introduced a model with magnetic levitation bearings, aimed at eliminating any possibility of oil contamination. The bearings were complex and expensive. Several other manufacturers have adopted the vertical arrangement especially for the smaller lower speed pumps. The design has been criticized because, without a bearing at the top end of the shaft, the mechanism is vulnerable to mechanical shock. Pfeiffer have overcome the problem by introducing a permanent magnet bearing on the high vacuum end of the shaft.

Although the turbomolecular pump offers higher pumping speed than the molecular drag pump, the latter is capable of attaining higher compression ratios. This fact has been exploited by CIT-Alcatel in a hybrid pump which combines the two structures. The design described by Maurice[17] is illustrated diagrammatically in *Figure 3.9*. At the inlet to the pump is a four-stage turbomolecular structure of 20 cm diameter. Beyond this on the same shaft is a cylindrical type of molecular drag pump. The drum diameter in this stage is 14 cm and 8 cm long and the cylindrical drum housing has five parallel helical grooves which decrease in depth towards the outlet. The whole assembly rotates at 24 000 rev/min to give a pump speed of $1620\,m^3h^{-1}$ with a

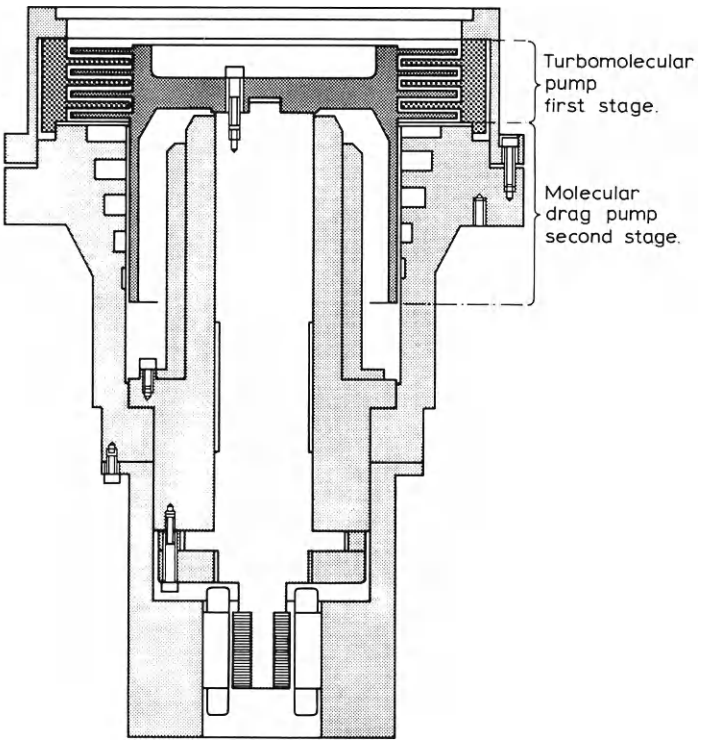


Figure 3.9 Sectional diagram of turbomolecular pump with a second stage molecular drag pump according to Maurice[17]

compression ratio for nitrogen of 10^{11}. With this large compression ratio it is possible to pump down to 10^{-6} Pa with the outlet at atmospheric pressure. However, to ensure molecular flow conditions at the inlet initially, it is necessary to back the pump out to 10 Pa on 'start-up'.

A novel feature of the CIT-Alcatel pump is the compressed air bearings, which eliminate any possible oil contamination and are claimed to give no wear. It does, however, need a compressed air storage tank to maintain the pressure whilst the rotor is slowing down in the event of a power failure. A later design also allowed the rotor to be driven by a compressed air turbine.

The turbomolecular pump, then, offers a high-speed, oil-free pump capable of evacuating to 10^{-8} Pa. Normally the relevant part of the pump can be baked to 100°C or so and there is no requirement for baffles or traps. It is, however, a precision engineered device running at high speed. Consequently it is expensive and in general requires regular servicing. In particular there is the problem of wear of the bearings and it is very susceptible to damage should small particles, such as chips of glass, become drawn into the structure. If it is

backed by a rotary pump, precautions should be taken to exclude oil-vapour, although the problem is less important than with a diffusion pump because of the high compression ratio for the heavier molecules.

Turbomolecular pumps are gaining in popularity especially for applications such as electron microscopes, where a fairly high throughput is required with complete freedom from oil contamination.

3.4 Sorption pumps

Sorption pumps, and indeed the pumps described in the following sections, depend on the adsorption and desorption processes described in Chapter 1 (Section 1.5.4). In that chapter it was shown that at any surface there are continuous processes of adsorption and desorption of the ambient gas taking place as a result of the bombardment of the surface by the gas molecules and the physical or chemical binding forces active between them and the molecules of the solid. Under equilibrium conditions the adsorption rate, v_a, and desorption rate, v_d, are equal and there is no net loss or gain of gas on the surface. However, if the equilibrium conditions are disturbed, by changing the ambient temperature or gas pressure for example, gas will be desorbed or adsorbed and the surface will act either as a gas source or a pump according to the direction in which the equilibrium has been changed.

From Equation (1.44) we saw that the rate of desorption was given by

$$v_d = BN_A \exp(-E/RT_s) \tag{3.2}$$

where B is a constant, N_A is the number of gas molecules already on unit area of the surface, T_s is the surface temperature and E is the binding or activation energy associated with the intermolecular forces. The rate of adsorption was given by Equation (1.45)

$$v_a = \frac{fP}{\sqrt{2\pi mkT_g}} \tag{3.3}$$

where f is the fraction of impinging molecules which will stick to the surface as a result of van der Waals or chemical bonding forces, and T_g the gas temperature. Thus, combining Equation (3.2) and Equation (3.3) gives the net rate of change of surface gas per unit area dN_A/dt as

$$\frac{dN_A}{dt} = v_a - v_d = \frac{fP}{\sqrt{2\pi mkT_g}} - BN_A \exp(-E/RT_s) \tag{3.4}$$

In order to use the surface or part of the surface as a pump, v_a must be greater than v_d and thus the principle of employing the sorption process for pumping is to arrange for the most favourable relationship between v_a and v_d. This is achieved by using surfaces with high values of E or f, for example, and/or by reducing the surface temperature T_s. The total pump speed is given by $A\,dN_A/dt$ where A is the effective surface area, a further parameter which can be exploited in the sorption pumping process. It is worth noting, however, that as the gas is adsorbed, N_A and thus v_d increases, causing the pump speed dN_A/dt to decrease until eventually a new equilibrium is reached as $dN_A/dt \rightarrow 0$. Thus pumps based on sorption processes will tend to saturate

with time unless they can be continuously activated. The total amount of gas that can be adsorbed will depend on the sticking factor, the activation energy, the surface temperature and the surface area.

The sorption process is exploited in most of the alternative pumps to the diffusion pump, such as the sputter ion pump, sublimation pump and cryogenic pump, discussed later. However, the term 'sorption pump' is applied specifically to a pump in which the gas is physically sorbed in a porous material cooled to liquid nitrogen temperature and which is normally employed for pumping down from atmospheric pressure to about 1 Pa as a backing pump.

This type of pump dates back to the days of Dewar in the late 19th Century. In early pumps activated charcoal derived from coconut was employed, but the process was rather uncontrolled giving very inconsistent results. With the advent of the diffusion pump, sorption pumping was abandoned and was only revived in the 1950s when the need arose for a hydrocarbon-free backing pump for the newly developed ion pumps. The development of sorption pumps for this purpose came with the use of so-called 'molecular sieve' material. Molecular sieve is synthetically prepared zeolite (alumino silicates), such as sodium alumino silicate, $Na_2O, Al_2O_3\, nSiO_2, xH_2O$. The crystals have a large amount of water of crystallization but, unlike other crystals, the structure does not collapse when the water is driven off. Instead it leaves a cavity structure with controlled pore diameter. The pore diameter, which depends on the method of manufacture, ranges from 4–10 Å and compares with the diameter of gas molecules which are of the order of 3 Å. *Table 3.2* lists information on the pore diameters of molecular sieve available from Linde, compared with that for activated charcoal. The material is in the form of small crystallites which

TABLE 3.2. Molecular sieve material available from Linde Division of Union Carbide Corporation

Type no.	*Material*	*Pore diameter* Å
4A	Sodium alumino silicate	4
5A	Calcium alumino silicate	5
13X	Sodium alumino silicate	10
–	Activated charcoal	50

are bonded into pellets of a few millimetres diameter with china clay or similar material. The porous nature of the crystallites gives a large effective area for adsorption. Measurements by the BET* method give the effective surface area as 600 m^2g^{-1}.

In practice, the difference in performance between the types of sieve is marginal, and the choice tends to be rather arbitrary; some workers prefer 5A whilst others recommend 13X.

Essentially the pump consists of a chamber containing the adsorbent material. It is usually connected to the pumping system by a valve as illustrated in *Figure 3.10*. The pump is 'activated' by baking to around 200°C with the inlet valve open and the valve to the vacuum system closed. Baking

* Brunauer–Emmett–Teller method, see Dushman[18] page 395

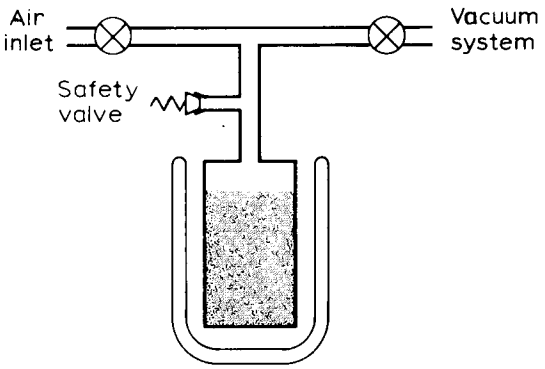


Figure 3.10 Arrangement for sorption pumping

drives off water vapour which constitutes the bulk of the gas adsorbed at room temperature. After baking, the inlet valve is closed and the pump refrigerated in a liquid nitrogen dewar. The valve to the system is then opened and the system will pump down to around 1 Pa. If it is acting as the backing pump to an ion pump it can be valved off once the ion pump is in operation. Since the amount of gas adsorbed can be several times the pump volume at STP, a pressure build-up will occur in a valved-off pump if the refrigerant is removed. A safety valve is therefore provided to cope with such an eventuality.

The design and performance of the pump are mainly dictated by three factors, namely: (1) the saturation effect; (2) the preferential pumping of the various gases, and (3) the thermal conductivity of the pellets. We can see the importance of the first two factors by considering the adsorption processes as expressed in Equation (3.4).

The total amount of gas which can be adsorbed per unit area $N_{A(sat)}$ can be ascertained from the equation by considering the equilibrium condition of $dN_A/dt = 0$.

Then $v_a = v_d$ and

$$N_{A(sat)} = \frac{fP}{B\sqrt{2\pi mkT_g}} \exp(E/RT_s) \tag{3.5}$$

Thus $N_{A(sat)}$ increases linearly with pressure and decreases exponentially with increasing surface temperature when E is positive, which is the case for physical adsorption. A plot of $N_{A(sat)}$ against P at constant temperature is called an adsorption isotherm and gives a useful assessment of an adsorbent material's performance.

The value of E for physical adsorption is almost independent of the adsorbent material but is very dependent on the gas species. It is equal to the heat of adsorption and increases as the boiling point increases. For example, helium with a boiling point of 4.2 K has a heat adsorption of about 590 J mol^{-1}. Hydrogen, with a boiling point of 20.4 K, has a value of 6.3 kJ mol^{-1}, whilst argon, nitrogen, oxygen and carbon monoxide, with boiling points around 77–90 K, have heats of adsorption of 12–17 kJ mol^{-1}.

The effect of this variation in E is illustrated in *Figure 3.11*, where adsorption isotherms for neon, helium and nitrogen obtained by Turner and Feinleib[19]

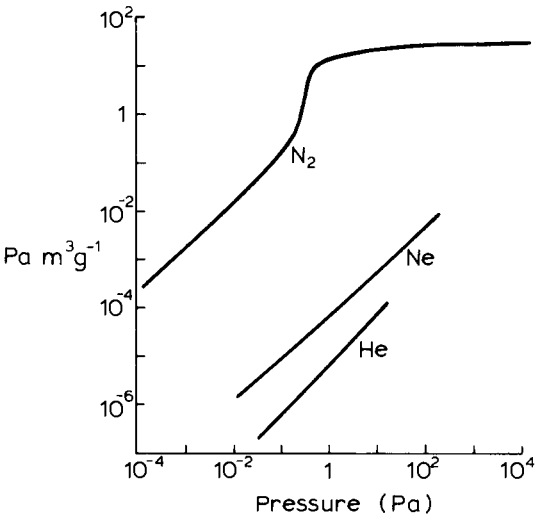


Figure 3.11 Adsorption isotherms for nitrogen, helium and neon on molecular sieve 5 A at 78 K, from Turner and Feinleib[19]

for molecular sieve are represented. The curves were obtained by evacuating the pump and system to about 10^{-4} Pa, admitting metered quantities of gas and measuring the equilibrium pressures. The curves show that at liquid nitrogen temperature, 78 K, much larger quantities of nitrogen are adsorbed than neon or helium, as expected from their activation energies. The saturation effect for nitrogen is in accordance with the suggestion that a complete monolayer is formed on the surface with 10 Pa m^2g^{-1} and that thereafter the adsorption is much reduced. 10 Pa m^3 of gas contains approximately 3×10^{21} molecules, which, divided by the surface area of a gram of sieve, i.e. 600 m^2, gives 5×10^4 molecules cm^{-2}. The latter is in general accordance with the molecular density for a monolayer.

From the isotherms, the performance of the sieve as a pump can be deduced in terms of ultimate pressure reached when a volume, V, is pumped from a pressure, P, by n grams of sieve material. As example, Turner and Feinleib have considered the pumping of air from atmospheric pressure by 5A molecular sieve chilled to liquid nitrogen temperature. *Figure 3.12* shows the ultimate pressure as a function of V/W, deduced from the isotherms, where V is the volume of the system and W the weight of sieve material. The curves show that for V/W values below 10^{-4} m^3g^{-1}, the partial pressure of Ne, He and N_2 are similar and that over the V/W range of 10^{-6} to 10^{-4} m^3g^{-1} the residual nitrogen pressure varies by less than an order of magnitude. In practice, neon and helium are adsorbed to a lesser extent. This is probably due to the large adsorption of nitrogen affecting the isotherms of the other gases, a factor not taken into account in *Figure 3.12*. Nevertheless, the curves are in general agreement with experimental data and some typical mass spectrometer analyses of residual gases after sorption pumping air are given in *Table 3.3*. The data are also taken from Turner and Feinleib's paper.

Thus a molecular sieve pump, pumping down from atmospheric pressure,

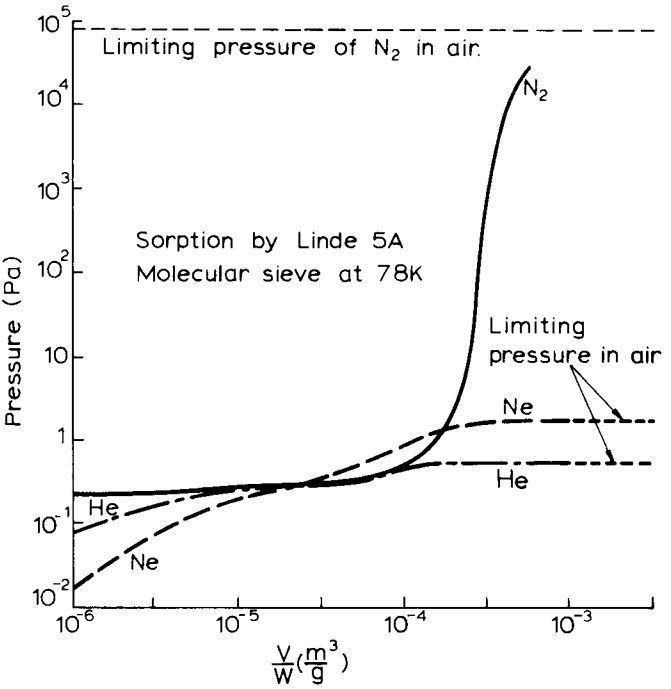


Figure 3.12 Predicted equilibrium pressure for helium, neon and nitrogen in air as a function of the pump load, from Turner and Feinleib[19]

can be expected to reach 1 Pa, the limitation being mainly due to the neon in the atmosphere. The amount of sieve required is not critical provided that V/W is less than 10^{-4} m^3g^{-1}. However, there are two further points to be taken into account in the design of a good pump. One is to ensure that the sieve pellets are adequately cooled, and the other is to ensure that the pump has sufficient conductance for the gas to reach all the sieve material. The sieve material has a low heat conductivity and in a badly designed pump the pellets remote from the cooled surface can remain warmer and act as a source of gas, so limiting the ultimate pressure. Similarly the pressure can be limited if the gas to be pumped cannot reach the pellets remote from the orifice. The pump design, therefore, aims at presenting the maximum surface area for cooling for the volume of adsorbant required but at the same time allowing adequate conductance for all the adsorbant to be effective.

Small pumps are often in the form of long tubes, whilst larger pumps will be annular in shape or consist of several tubes in parallel. Some pumps contain a central gauze cylinder, restricting the sieve pellets to the area close to the wall and allowing a high conductance for pumping. *Figure 3.13* illustrates some of the designs adopted for sorption pumps.

So far the use of a single pump for evacuating down from atmospheric pressure has been considered. Assuming there is an adequate quantity of sieve, the ultimate pressure is then limited by the partial pressure of neon in the atmosphere to around 1 Pa. If however, the sorption pump is pre-evacuated a lower pressure can be obtained.

TABLE 3.3. Analysis of residual gases after sorption pumping of air with molecular sieve type 5A according to Turner and Feinleib[19]

Initial temp. = 293 K		*Cooling temp.* = 77 K			
Pump load					
Volume (m^3) / Weight of Mo. Sieve (g)		3×10^{-4}	2×10^{-5}	2×10^{-5}	2×10^{-6}
Pumping time (min)		24	10	25	10
Total pressure (Pa)		2.7×10^{3}	2.0	1.2	1.3
Partial pressures (Pa)	N_2	2.7×10^{3} (100%)	1.5×10^{-1} (7.3%)	2.0×10^{-1} (16.7%)	1.3×10^{-1} (10%)
	O_2	–	2.4×10^{-1} (12%)	1.3×10^{-1} (11.1%)	6.7×10^{-2} (5%)
	A	–	4.7×10^{-2} (2.3%)	4.0×10^{-2} (3.3%)	4×10^{-3} (1%)
	Ne	–	1.2 (60%)	6.4×10^{-1} (53.4%)	9.1×10^{-1} (68%)
	He	–	1.3×10^{-1} (6.7%)	8×10^{-2} (6.7%)	9.3×10^{-2} (7%)
	H_2O	–	2.0×10^{-1} (10%)	4×10^{-2} (3.3%)	1.3×10^{-1} (10%)
	H_2	–	8×10^{-3} (0.4%)	1.3×10^{-2} (1.1%)	3.3×10^{-3} (1%)
	CO_2	–	2.1×10^{-2} (1%)	5.3×10^{-2} (4.4%)	2.7×10^{-3} (1%)

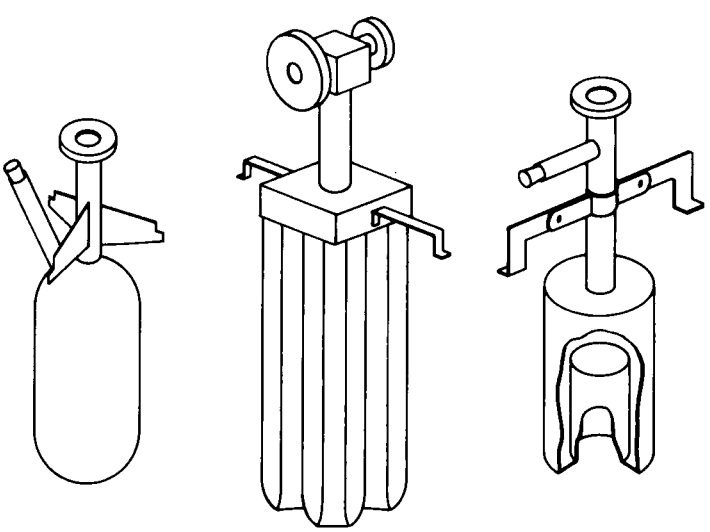


Figure 3.13 Some typical constructions of sorption pumps

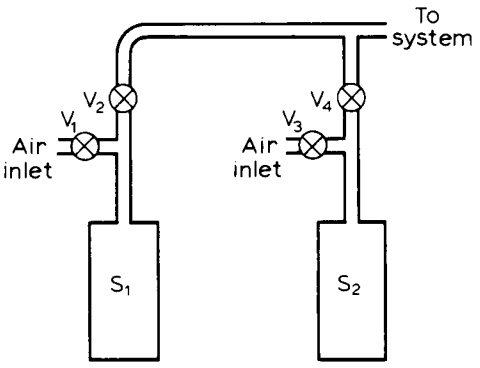


Figure 3.14 Connection for two sorption pumps working in tandem

Consider two sorption pumps used in tandem, *Figure 3.14*. The two pumps are activated by baking with valves V_1 and V_3 open and V_2 and V_4 closed. Valves V_1 and V_3 are then closed and V_2 and V_4 opened, and pump S_1 refrigerated. When the pressure flattens off at the lower pressure, valve V_2 is closed and pump S_2 refrigerated to continue the pumping. The performance of two pumps working in this way can be predicted from the appropriate isotherms. The first pump reduces the nitrogen pressure by 6 orders to around 10^{-1} Pa, but the neon pressure is only reduced by an order to about 5×10^{-1} Pa. The residual gas is now mainly neon and the second pump will only reduce this to around 5×10^{-2} Pa. The nitrogen pressure on the other hand will be below 10^{-5} Pa. Although the gain in total pressure is not very great, it is still preferable to use two pumps in tandem than a single pump of double the capacity. First, it allows the partial pressure of active gases, nitrogen, oxygen and water vapour to be considerably reduced. Secondly, whilst one pump is being used, the other can be regenerated to give a high capacity continuous pumping cycle. This sytem is therefore advocated for larger ultrahigh vacuum equipment.

Consider now the possibility of pre-evacuating the sorption pump by an alternative method, for example by a rotary pump. Then the neon partial pressure will be reduced by the same amount as the nitrogen before operating the sorption pump and, on refrigeration, reduction to high or even ultrahigh vacuum pressures becomes possible. It may be argued that one of the main attributes of a sorption pump is its freedom from contamination and if it has first to be evacuated by a rotary pump or similar pump this advantage is lost. However, as Read[20] has pointed out, this is not necessarily so. The molecular sieve is a very good adsorber of oil vapour and the pump itself can be used as an effective trap. Read designed a pump with this feature incorporated which is illustrated in *Figure 3.15*. By evacuating the sorption pump with a diffusion–rotary pump combination to 10^{-7} Pa, he attained a pressure of 10^{-9} Pa when it was refrigerated to 78 K.

For high vacuum performance better thermal contact between the cooled surface and the sieve material is required. When pumping down from atmospheric pressure, the heat from the pellets remote from the cooled walls is conducted away mainly by the high-pressure gas. When the pump is evacuated

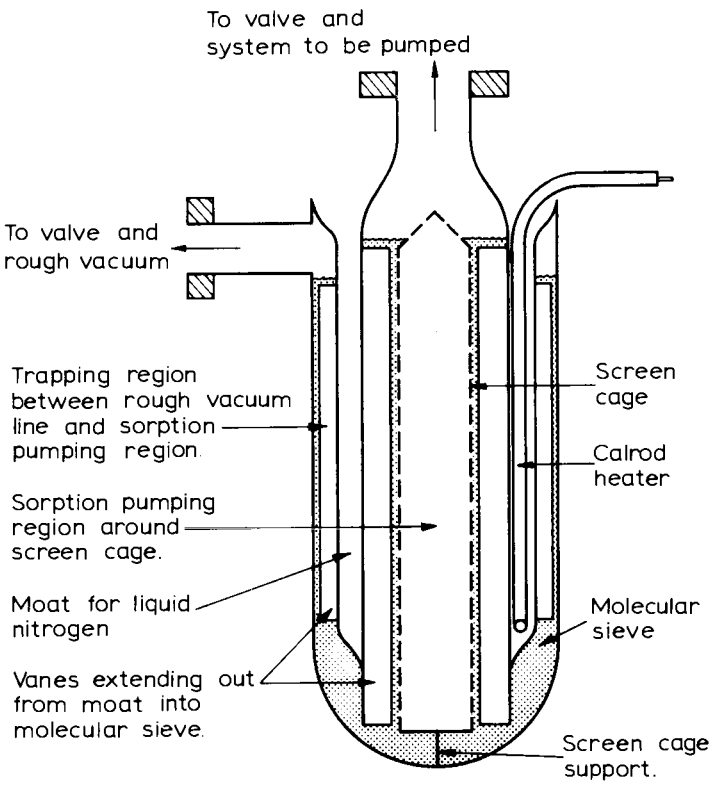


Figure 3.15 Sectional diagram of sorption pump for high vacuum pumping according to Read[20]

this no longer applies and one has to rely on the poor thermal conductivity of the sieve material. The effect of this is illustrated in *Figure 3.16* where the temperature measured at the centre of a 40 mm diameter tube filled with 13X pellets is plotted against time of refrigeration. It is seen that in the case of the pre-evacuated run the temperature fell very slowly and even after refrigerating overnight the temperature had not gone below 190 K. The commercially available pumps, designed for pumping from atmospheric pressure, are normally unsuitable for pumping to high or ultrahigh vacuum pressures. A more elaborate design is required with some method of improving the thermal conduction. A number of designs have been described in the literature[20,21,22] where the sieve material is restricted to a thin layer not more than a few millimetres thick on the cooled surface, and/or extra cooling fins or similar arrangements are made to increase the refrigerated surface area. The sieve can be held in position by a gauze, or in one case the sieve was bonded on to the metal with a suitable adhesive.

Operation at pressures lower than 1 Pa has two main applications. First, as a backing pump for ion pumps. If the backing pressure can be reduced to 10^{-3} or 10^{-4} Pa then an ion pump can be started up with a simpler and cheaper power supply (see Section 3.7). Secondly, as a booster pump to deal with outgassing or other sudden influx of gas in high or ultrahigh vacuum systems.

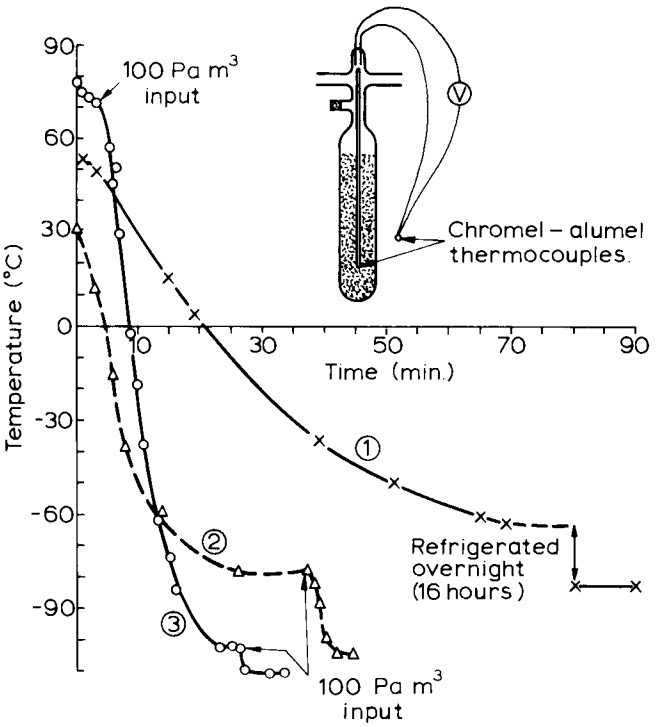


Figure 3.16 Temperature fall with time at the centre of a zeolite pump on immersion in liquid nitrogen, showing the effect of gas pressure. 1, Pump pre-evacuated before refrigeration; 2, Pump at atmospheric pressure when refrigerated; 3, Pump pre-evacuated, 100 Pa m^3 let in after 5 min refrigeration

The attributes of the sorption pump for this application are its large gas capacity and freedom from contamination.

A problem which confronts the user of sorption pumps should be mentioned. It concerns the physical structure of the sieve material. Continuous use of a sorption pump tends to powder the adsorbant and it is difficult to prevent the powder entering the vacuum system, where it can form an undesirable deposit or foul metal valve seatings. Most pumps will contain a gauze or porous plug at the orifice to minimize the problem, but the system should be designed to ensure that there are no metal bakable valves close to the pumps. As a further precaution it is wise to recharge the pump with fresh molecular sieve at intervals when powdering appears to be occurring.

3.5 Cryogenic pumps

It was shown in the last section that a refrigerated surface will pump gas by physical adsorption whilst non-equilibrium conditions prevail. The gas forms a monolayer on the surface but thereafter the pressure rises until equilibrium conditions are established (cf. *Figure 3.11*). In the equilibrium condition, with

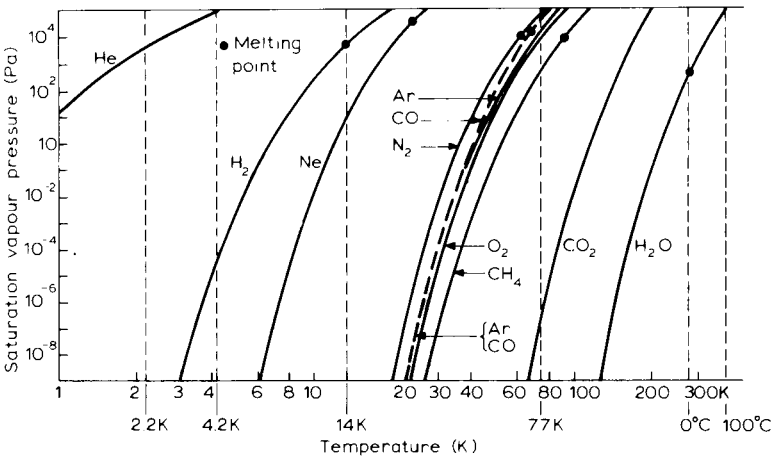


Figure 3.17 Saturation pressure as a function of temperature for gases commonly encountered in vacuum systems

the gas temperature equal to the surface temperature of the walls, $v_a = v_d$. The final pressure reached will be the saturation vapour pressure of the gas concerned. Attempts to raise the pressure above the saturation pressure result in the gas condensing on the surface and building up a multilayer, as a solid or liquid. The saturation vapour pressure is very dependent on temperature and lowering of the temperature offers a method of reducing the pressure within a vacuum chamber. The effect is illustrated in *Figure 3.17* where the saturation pressures for the gases commonly encountered in vacuum systems, are plotted as a function of temperature. At liquid nitrogen temperature, 77 K, the pressure of water vapour alone is significantly reduced, to below 10^{-10} Pa. The saturation pressures of nitrogen, oxygen, carbon monoxide, etc. are still around atmospheric pressure. If the temperature is lowered to liquid hydrogen temperature, 20 K, then the above-named gases have saturated pressures in the ultrahigh vacuum region, below 10^{-6} Pa. Neon, hydrogen and helium are the only gases still having high saturation pressures. At liquid helium temperature 4 K the neon pressure is below 10^{-10} Pa and the hydrogen pressure is around 10^{-5} Pa. Thus, apart from helium and hydrogen it is possible to reduce the pressure in a vacuum system to ultrahigh vacuum by cooling the system to liquid helium temperature. If a lower temperature, say 2.2 K is achieved, helium alone will remain above 10^{-8} Pa. The pumping of gases by condensation is known as cryopumping.

Refrigerating the whole vacuum system to liquid helium temperature is not very practical and cryopumping is therefore carried out by cooling only part of the surface, or by using a special refrigerated appendage as a cryogenic pump or cryopump. The pumping speed and ultimate pressure of such a pump can be deduced from Equation (3.4). Since, however, at equilibrium $v_a = v_d$, v_d can be replaced by $v_{a(T_s)} = f_{(s-s)} P_{sat} / \sqrt{2\pi m k T_s}$, where $f_{(s-s)}$ is the sticking factor for gas at temperature T_s on a surface at the same temperature and P_{sat} is the saturation vapour pressure at temperature T_s. Rewriting Equation (3.4) we can then obtain the rate at which gas is adsorbed on unit area of refrigerated surface as

$$\frac{dN_A}{dt} = \nu_{a(T_g)} - \nu_{a(T_s)}$$

$$= \frac{1}{\sqrt{(2\pi mk)}} \{ f_{(g-s)} P T_g^{-1/2} - f_{(s-s)} P_{sat} T_s^{-1/2} \} \tag{3.6}$$

where $f_{(g-s)}$ signifies the sticking factor for gas at temperature T_g on a surface at temperature T_s.

The ultimate pressure is reached when $dN_A/dt \rightarrow 0$ then

$$P_{ult} = P_{sat} \frac{f_{(s-s)}}{f_{(g-s)}} \left(\frac{T_g}{T_s} \right)^{1/2} \tag{3.7}$$

Thus, because only a part of the vacuum system is at temperature T_s, the ultimate pressure is somewhat higher than the saturation pressure corresponding to T_s.

Substituting for P_{sat} in Equation (3.6) gives

$$\frac{dN_A}{dt} = \frac{f_{(g-s)}}{\sqrt{2\pi mk T_g}} (P - P_{ult}) \tag{3.8}$$

$f_{(g-s)}$ can be high for gas molecules incident on a very low-temperature surface, measurements give values from 0.8 to unity. Therefore, for pressures more than an order higher than the ultimate, the pumping action of the surface approaches closely the theoretical limit dictated by the number of molecules incident on the surface per second (cf. Equation (3.3)). This is considerably better than can be obtained by a diffusion pump or turbomolecular pump and, unlike the sorption pump, the surface does not saturate under high gas loads.

Having said this, it must be pointed out that if the layer on the condensation surface becomes too thick it can flake off and, in any case, there will be a temperature gradient across it which will cause the pumping action to deteriorate appreciably. In general, therefore, because of the limited capacity, cryopumping is combined with an alternative pump to evacuate the system down to 10 Pa or less. Nevertheless, cryopumping offers a high-speed, clean pumping system capable of dealing with fairly high gas loads down to pressures of 10^{-8} Pa or better, provided helium can be removed. Also it can be easily regenerated by simply allowing the temperature to rise and pumping away the desorbed gas with a roughing pump.

Cryopumps fall basically into two categories, those for large systems, such as space simulation chambers and particle accelerators for nuclear fusion experiments, and those for small systems of up to say, a cubic metre where freedom from any contamination is essential.

For the large systems the traditional approach of cooling the cryopanel with liquid helium is normally employed. In general the pumps in this category are specially designed to fit the system and make use of as much of the surface area as possible for maximum pump speed. A number of such designs have been described in the literature[23,24,25]. The basic designs are similar. Since the pumping speed is limited by the number of molecules incident on the projected surface area or the orifice, there is no gain in having a porous or roughened surface, so that a simple flat metal panel suffices. Liquid nitrogen cooled screens, permeable for gas but impermeable to radiation, usually surround the panel to reduce the dissipation and thus conserve the liquid helium. Also the

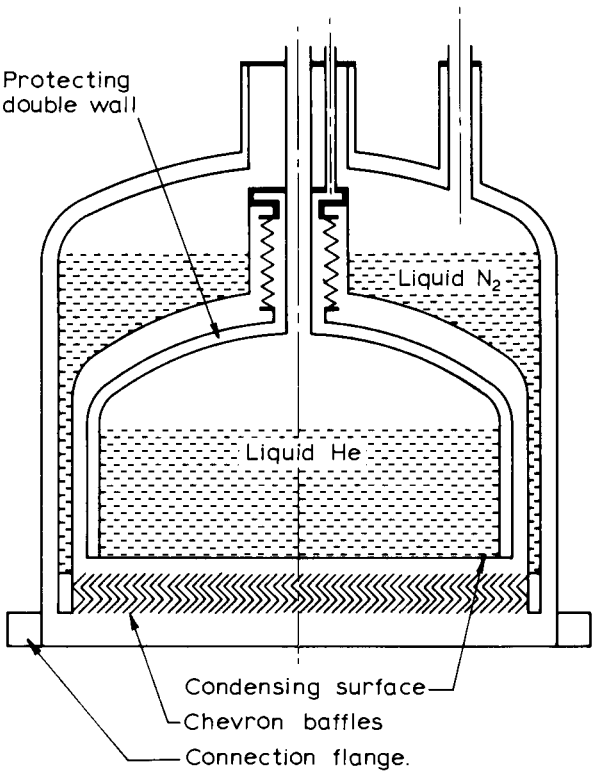


Figure 3.18 Liquid helium cooled cryopump according to Benvenuti[24]

heat shield will cryopump some of the gases such as CO_2 and H_2O. The pump described by Benvenuti[24] is fairly typical of designs for attaching as appendage pumps to the vacuum system; a similar design had been described by Hengevoss *et al.*[26]

The pump is shown diagrammatically in *Figure 3.18*. Basically it consists of a liquid helium container with a cold surface, surrounded by a liquid nitrogen container holding the cold chevron baffle. The cold surface is plated with silver to give a low thermal emissivity. The pump was made in two sizes, 800 mm diameter with a liquid helium capacity of 75 l and 320 mm diameter with a liquid helium capacity of 10 l. When using a single chevron baffle the hydrogen pump speed for the larger pump was 30 m^3s^{-1} and for the smaller pump 4.5 m^3s^{-1}. The pump speed for other gases was lower. The ultimate pressures obtained were in the region of 10^{-11} Pa.

An important factor with such pumps is the liquid helium consumption. For the large pump, 5–10 l were consumed in the initial cooling period during filling and thereafter about 1 l per day was lost. For the smaller pump the figures were 1–2 l and 0.25 l per day respectively. A later modification of the design has reduced the liquid He consumption by a factor of five.

Normally the pumps are designed to be part of the vacuum chamber with a sizeable proportion of the chamber walls cooled to liquid nitrogen tempera-

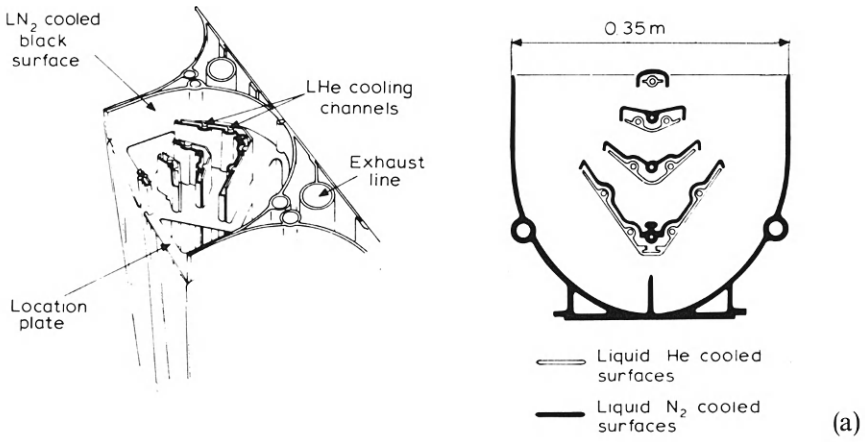
ture which would allow ultrahigh vacuum to be obtained without a baking cycle. Some designs incorporate a helium liquefying plant with the helium circulating through the chamber in a closed loop[27,28] Although the chevron baffle offers the most effective heat shield, it seriously reduces the pump speed of the cryopanel. Other designs have been introduced which have a more open structure, giving a higher pumping speed at the expense of a higher consumption of liquid helium. One such design is that developed for the neutral injection boxes of the Joint European Torus (JET) at Culham Laboratory, described in a review article by Hands[29]. The pump is of modular construction, each module, 6 m long and 0.35 m across, contains four cryopanels with liquid nitrogen-cooled shields. The cross sectional configuration is illustrated in *Figure 3.19(a)*. Ten modules mounted side by side form one of two units employed for each neutral injection box (*Figure 3.19(b)*). The pump speed of the unit for hydrogen is around 5×10^3 m^3s^{-1}, requiring about 40 W of refrigeration power for the helium panels and 10 kW for the liquid nitrogen panels.

For more modest vacuum systems the cost of designing an integral pump and of the liquid helium consumption, makes this type of design unattractive. However, the development over the last 20 years of small refrigerating machines with a few watts of cooling power at 20 K, based on regenerative isentropic expansion cycles using helium gas, has changed the situation and made cryopumping a practical proposition for many other applications.

An early system of this type was described by David and Venema[30] using a two-stage refrigerator based on the Stirling cycle. The first stage, giving a temperature of 50–80 K, was used to cool a heat shield, whilst the second stage of 20 K was connected to the cryopanel. The refrigerator they used was relatively clumsy, producing a considerable amount of noise and vibration. Since then there have been significant advancements in the design of small cryogenerators. The Stirling cryogenerator has been improved to give a more compact and quieter machine and design modifications have been made to improve the performance. The most important contribution has been that of Gifford and McMahon[31] whose design has been mainly responsible for the commercial success of these small cryogenerators.

The principle of the Gifford–McMahon cryogenerator is illustrated in *Figure 3.20*. It consists of three basic components: the compressor, the regenerator (heat exchanger) and the displacer. At the start of the cycle, compressed gas is let into the displacer cylinder via valve V_1, the displacer being at the bottom of its stroke and V_2 closed. V_1 is then closed and the piston raised so that the compressed gas passes through the regenerator, where it is cooled, into the lower parts of the displacer cylinder. V_2 is then opened expanding the gas and thus cooling it further. The piston is then pushed down to the bottom of the cylinder, the expanded gas being pumped by the compressor, V_2 is closed and the cycle repeated. In the practical system the regenerator will normally be incorporated into the displacer piston and the piston can be actuated pneumatically by the working gas itself. Also, a second stage is mounted on the same cooling head to give a lower temperature and allows the first stage to be used for the heat shield.

The main difference between the Gifford–McMahon and the Stirling cryogenerators is that in the former the compressor is separated from the cooling head, whereas in the Stirling cryogenerator, compression of the gas is



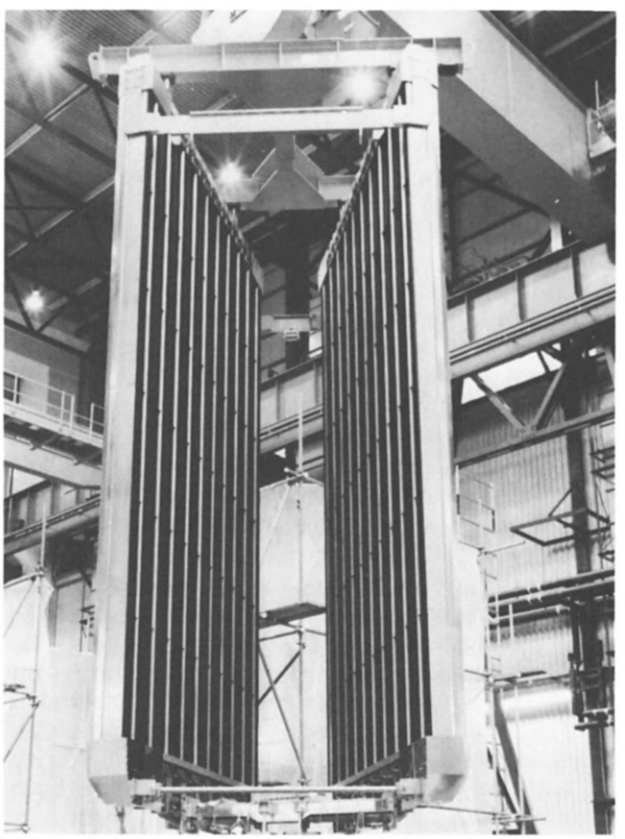


Figure 3.19 Cryogenic pump designed for the central injection boxes of JET. (a) Schematic of the pump elements; (b) Photograph of a completely assembled cryopump system. (Courtesy JET Joint Undertaking)

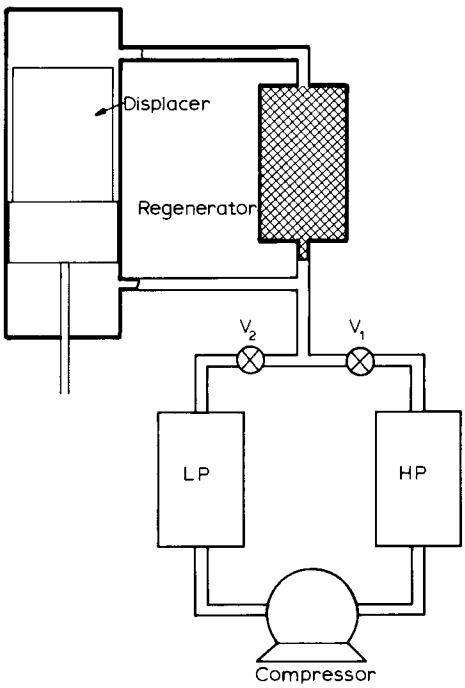


Figure 3.20 Principle of a single stage Gifford–McMahon cryogenerator

obtained by a piston moving out of phase with the displacer. Although the Stirling cryogenerator may be more efficient, separating the compressor from the cooling head allows the cooling head to be more compact and the main source of vibration to be isolated from the vacuum system. A typical configuration for a small cryopump is illustrated in *Figure 3.21* taken from a review article by Bentley[32] to which the reader is referred for more details. The first stage cools the heat shield to around 77 K and the second stage cools the cryopanel to 20 K, although the temperature attained will depend on the 'heat load' placed upon it. From *Figure 3.17* it is seen that neon, hydrogen and helium will not be pumped significantly at 20 K and so a porous surface, such as activated charcoal is normally coated on to the back surface of the cryopanel to sorption pump these gases effectively.

In summing up we can say that cryopumping offers distinct advantages over other pumping methods for achieving ultrahigh vacuum. The method is completely clean, capable of low ultimate pressures with high pumping speeds. For large systems which are already expensive in themselves, cryopumping is a very attractive approach, with the pump designed to fit the system. Such pumps use liquid helium and require a fairly elaborate gas recovery system. The development of small regenerative cycle cryogenerators now extends the application of cryopumping to much smaller systems with commercial pumps which are simple to install and to operate.

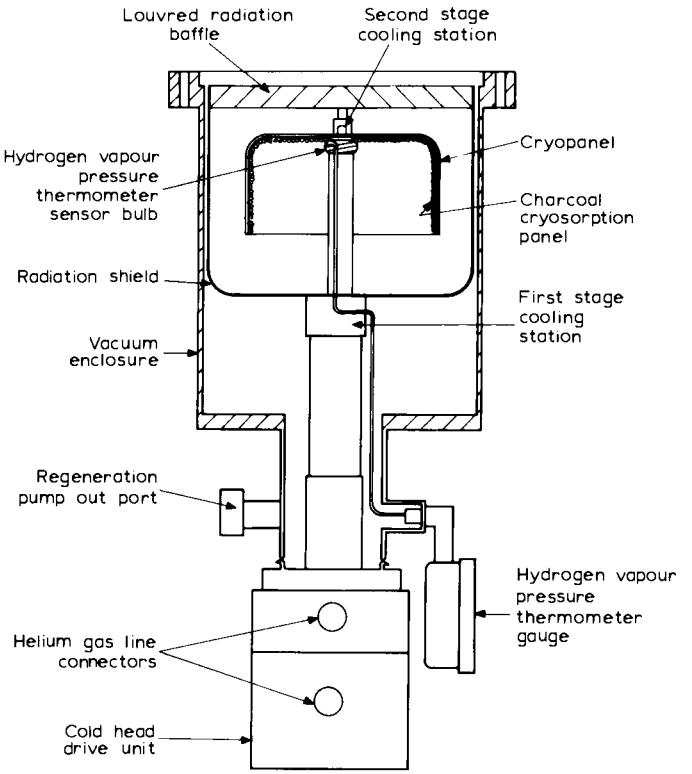


Figure 3.21 Typical small cryogenerator pump

3.6 Sublimation and getter pumps

The previous two sections describe pumps which depend on the physical adsorption of gases on refrigerated surfaces for the pumping mechanism. This section deals with pumps which rely on chemical sorption processes for trapping the gas.

Chemisorption is characterized by a much stronger bond between the gas layer and the surface as a result of a chemical or pseudo-chemical reaction which takes place between the gas and surface molecules. This is reflected in the binding energy E in Equation (3.2). For example, the value of E for oxygen chemisorbed on titanium is 1000 kJ mol^{-1} whereas for oxygen physically adsorbed on metal surfaces it is only 12–17 kJ mol^{-1}. In consequence the desorption rate v_d for chemisorbed gases is low, even at room temperature. Indeed for some chemisorption processes the reaction of gas with the surface occurs only at room or elevated temperature and reducing the surface temperature will prevent or slow down the reaction, so that the surface becomes less effective as a trap for gas.

Referring to Equation (3.4), we find that for chemisorption a favourable

relationship between v_a and v_d can be obtained without refrigeration, provided the chemical reaction can take place between the surface and the gas. Generally the bonding occurs only between the first few layers of gas and the surface and thereafter the gas is hardly adsorbed. In this respect the chemisorption surface acts in a similar way to a zeolite sorption pump and either a large surface area is required which can be activated or the active surface must be continuously replenished. Because of the strong bonding, it cannot normally be regenerated by raising to an elevated temperature. In some cases the gas will diffuse into the sorbant and react throughout the bulk, but in any case the capacity will be limited and pumps using chemical sorption are only used when the pressure has already been reduced to at least 1 Pa. Infact their main application is as an auxiliary booster pump providing a high pumping speed for a limited period.

The chemisorption of gas molecules on a surface is often referred to as gettering and has been the subject of extensive study since the first use of red phosphorus to improve the life of filament lamps. Today all electronic tubes, such as CRTs, will contain a 'getter' to improve the vacuum after sealing from the pump and to maintain the vacuum during life. Almost any metal that reacts readily with the active gases such as oxygen, nitrogen, hydrogen and carbon dioxide can be employed as a getter, but those with the best performance have been selected and perfected over the years. Two types of getter have evolved from these studies, evaporated getters whereby a metal film is evaporated on to the tube walls, and bulk getters which are activated by heating but are not evaporated. Commonly the evaporated getter film operates at room temperature whilst the bulk getter operates at a temperature above ambient. In the same way there are two types of chemisorption pumps; those in which a metal film is evaporated on to the pump walls, either continuously or at intervals, the so-called sublimation pump, and the more recently introduced pump in which non-evaporated getters are used.

The getter material used for the pump depends on a number of criteria, the most important being the chemical activity. Metals commonly used for gettering are listed in *Table 3.4*. Also given are values of adsorption capacity, although these should only be taken as indications of the getter performance as there is a wide variation in the reported values. The way in which the films are deposited, the temperature during evaporation and in subsequent operation and even the way in which the measurements are carried out affect the results.

For electronic tubes the most popular evaporated getter is barium and it is the metal which has been most extensively studied for gas 'clean-up'. Various non-evaporated getters have been employed, but at the present time an aluminium–zirconium alloy is favoured. Barium is chosen not only for its chemical activity but also because it can be readily volatized from a stable alloy or from a metal capsule holding the getter material. For electronic tubes a single 'flash' evaporation is desirable and the barium getter is especially suitable for this action. Although pumps using barium have been described in the literature with a heated crucible of barium[33], it is not easily adaptable to continuous or repetitive operation and has not found commercial application as a sublimation pump. The refractory metals such as molybdenum, tantalum, zirconium and titanium are far more suitable since they can be run as a heated filament or ribbon, continuously evaporating for some time before they burn

TABLE 3.4. Adsorption capacity of various metals used for gettering for the active gases expressed as $Pa\,m^3g^{-1}$ ($Pa\,m^3m^{-2}$)

	CO		CO_2		H_2		N_2		O_2	
Barium film	0.95	(0.95)	0.8	(0.8)	6.0	(6.0)	0.43	(0.43)	6.6	
Cerium bulk (Misc. metal)	–	–	0.3 → 6.0		6 → 8.5		0.43 → 2.1		2.8 → 6.7	
Molybdenum film		(0.20)	–	–		(0.07)		(0.07)		(0.7)
Tantalum film		(0.10)	–	–		(0.7)		(0.13)		(1.3)
Titanium bulk	–	–	6.6	–	–	–	21.3	–	12.0	–
Titanium film		(0.19 → 0.86)		(0.013 → 0.93)		(0.29 → 8.6)		(0.01 → 0.45)		(0.93)
Zirconium bulk	0.48 at 800°C		0.4 at 800°C		1.8 at 350°C		0.2 at 800°C		0.26 at 400°C	

out. All these metals have been advocated as the getter material in sublimation pumps but titanium is generally recognized as being the best material and is the one that is exploited in commercially available pumps.

The design of a titanium sublimation pump is very simple. Normally it will consist of a stainless steel cylinder with a flange at either end. One flange connects the cylinder to the vacuum system, whilst the other flange carries the electrical leadthroughs and supports for the filament, which is positioned axially in the cylinder. Alternatively, the flange with the filament can be mounted directly on to the vacuum chamber where the titanium will evaporate on to the chamber walls or on to a suitable screen mounted around the filament. A typical construction is illustrated in *Figure 3.22*. More than one filament is normally mounted in the pump to allow a long operational life between filament replacement. The cylinder may be water cooled or in some designs liquid nitrogen cooled.

The evaporation rate is an exponential function of the temperature and to obtain a reasonable evaporation rate the filament has to be run near the melting point of the titanium. As a result, slight non-uniformity of the filament can lead to rapid burn out. Two methods can be employed to overcome this problem. The filament can be wound on a supporting mandrel of higher melting point metal, or the titanium can be alloyed with another metal to raise the melting point. Clausing[34] described two such techniques for sublimation pumps at the time of their inception. For the supported filament, he wound a tantalum rod first with a layer of niobium wire and then with two layers of titanium wire. When heated, the niobium tended to alloy with the titanium

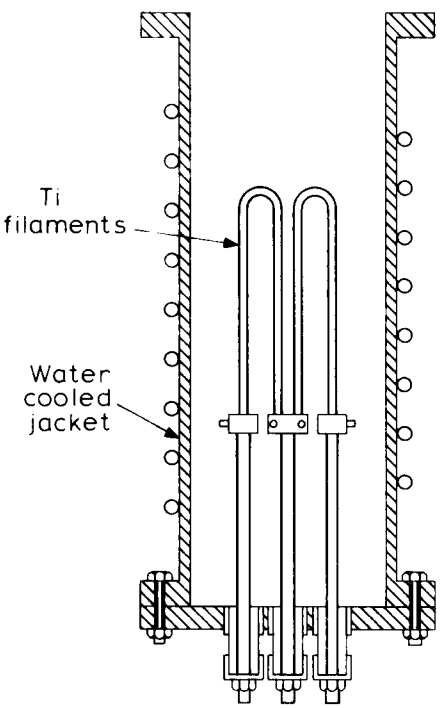


Figure 3.22 Construction of a titanium sublimation pump

raising its melting point. For the other method he used a titanium–tantalum alloy. Today, both techniques are applied to commercial pumps, but for the supported filament the normal arrangement is to wind the titanium wire together with a molybdenum wire of larger diameter on to a tungsten mandrel. The larger diameter molybdenum wire tends to bind the titanium close to the mandrel to give a good thermal contact. The preferred alloy for the other method is 85% titanium + 15% molybdenum.

McCracken and Pashley[35] found that the alloy filament gave better reproducible evaporation rates, with peak rates up to 2×10^{-5} g cm^{-1} s^{-1} and that 40% of the available titanium could be evaporated. The main limitation was crystal growth, causing embrittlement, which occurred rapidly at higher temperatures.

Lawson and Woodward[36] have made a more detailed study of the performance of alloy filaments. They showed that only titanium was evaporated and that the large crystal growth occurred when the composition of the filament reached 74% Ti 26% Mo and did not depend on the evaporation temperature used to reach that point. The evaporation rate as a function of temperature is illustrated in *Figure 3.23* and shows good agreement

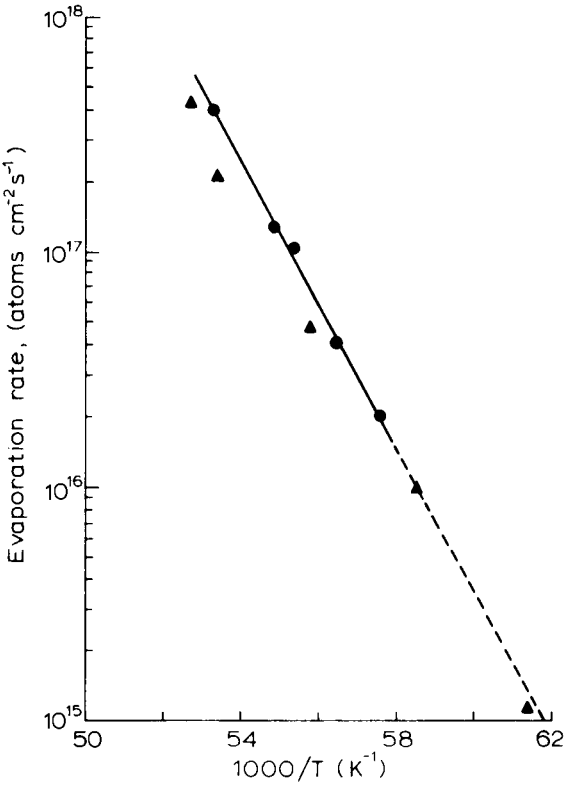


Figure 3.23 Evaporation rate of titanium from an 85% Ti/15% Mo alloy filament as a function of temperature. ● Lawson and Woodward[36] results; ▲ McCracken and Pashley[35] results

between the two investigations. McCracken and Pashley advocated running the filament under constant current conditions, the increased power as the filament diameter decreased compensating for the change in titanium content. However, Lawson and Woodward found that the evaporation rate increased with time under these conditions and that running at constant voltage gave a more consistent rate. Although the current falls as the filament thins, there is a net rise in temperature to compensate for the smaller surface area of the filament and the change in content. At the onset of crystal growth the evaporation rate decreased markedly and this can be considered as the end of its useful life.

Consider now the gas pumping. Clearly the rate of pumping is determined primarily by the area of uncombined titanium film available. From this it follows that the speed of pumping is a function of the rate of evaporation and that to maintain a given speed, the same amount of getter must be evaporated per second as is consumed by combination with the gas in the system. Since the amount consumed depends on pressure, it would not be very practical to design a pump for continuous pumping over a wide pressure range and the practical pumps are operated in pulses so as to give a satisfactory average pump speed. If the average rate of deposition is greater than the rate at which gas combines with the film then there will always be a fresh getter surface available and the speed will be a maximum. Thus the maximum will be approached asymptotically with increasing rate of deposition and thereafter increasing the deposition rate will hardly affect the pump speed. Thus the pump works most efficiently at the minimum evaporation rate required for maximum pump speed. The power supply is usually arranged to give a variable time on and off measured in minutes, for example an 'on' time of up to five minutes and an 'off' time of up to 60 minutes. The filament current can also be varied. The manufacturer recommends time intervals and currents for pumping at various pressures, typically at 10^{-4} Pa the evaporation would be repeated every five minutes whilst at 10^{-6} Pa a 30 minute 'off' interval is adequate, but it may be useful to adjust the supply to give optimum performance in a particular situation. In any case to economize on filament life it is porpitious to make adjustments to the supply as the system pumps down.

The 'non-evaporation' getter pump utilizes the zirconium–aluminium alloy getters developed for the electronic tube industry by SAES getters[37]. Alloying zirconium with up to 30% aluminium results in crystalline structures of intermetallic compounds Zr_5Al_3, Zr_3Al_2 and Zr_5Al_4. The alloy with maximum activation consists of the bi-phase structure containing Zr_5Al_3 and Zr_3Al_2 formed with approximately 15% by weight of aluminium, the optimum alloy composition depending slightly on the gas being adsorbed[38]. The sorption process is controlled by diffusion and therefore the sorption rate increases markedly with temperature, especially above 200°C. However, even at room temperature gases are adsorbed, especially hydrogen, and indeed above 400°C hydrogen can be released. The other active gases form stable compounds and their re-emission has not been observed up to 1000°C. Inert gases such as argon are not adsorbed. The getter is activated at about 800°C for several minutes in vacuum, when the adsorbed surface layers are diffused into the interior. The getter may be submitted to further activations after exposure to air if required until all the active material has been saturated.

Various constructions of a 'non-evaporation' getter pump are possible, the

main design criterion is the need to activate the getter and run it at a temperature around 400°C. Commercially available getters can be mounted in a bulb, activated by eddy current heating, and operated in an oven or heated by radiation. Della Porta and Ferrario[39] described a small appendage pump of $10^{-2}\,m^3s^{-1}$ which incorporated a Bayard–Alpert type gauge. The zirconium–aluminium alloy was coated on to both sides of a metal substrate which was then pleated to give a maximum active area of about 180 cm^2. The pleated getter 'cartridge' was mounted co-axially round the gauge structure, also enclosing four heaters used to activate the getter and maintain it at any desired operating temperature, *Figure 3.24*. The pumping speed for nitrogen was about $3 \times 10^{-3}\,m^3s^{-1}$ at room temperature, falling rapidly as the gas was adsorbed. At 370°C the pump speed was $9 \times 10^{-3}\,m^3s^{-1}$ falling to an equilibrium value around $6 \times 10^{-3}\,m^3s^{-1}$ on adsorbing $3 \times 10^{-3}\,Pa\,m^3$ of gas. The pump speed for hydrogen was about four times higher and was constant even at room temperature.

The gauge allows the pressure to be monitored, but similar cartridges can be used as pumps without the gauge insert. Tuck[40] described such a pump giving

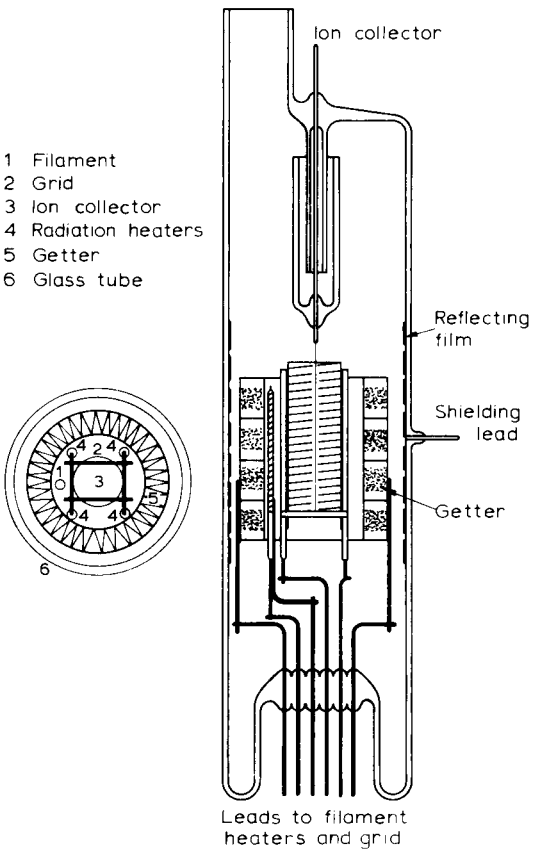


Figure 3.24 Getter pump using non-evaporation getter cartridge Della Porta and Ferrario[39]

a speed of 5×10^{-2} $m^3 s^{-1}$ used as an appendage pump in the processing of Klystron tubes.

Although the pumping speed of the Zr–Al getter pump is much less for a given pump size than for a titanium sublimation pump, it is easier to control and uses the getter material more efficiently. The cartridges can be changed when they are fully saturated but they will have a very much longer life than the filaments in a sublimation pump.

In general it can be said that sublimation and getter pumps offer useful additional pumps to deal with gas loads introduced into vacuum systems by, for example, outgassing. The difficulty of ensuring maximum activation and of controlling their pump speed reproducibly, does not make them very suitable as primary ultrahigh vacuum pumps. Also they will only pump active gases leaving the inert gases unpumped.

3.7 Ion pumps

It has been known for many years that ionized gases are more readily 'trapped' than neutral molecules, evident in the problem of gas clean-up in thyratrons and other low pressure gas discharge devices. The detrimental clean-up process, which is a major cause of tube failure, has prompted extensive investigation over the years, mainly aimed at reducing the effect. To the vacuum engineer it also appeared as an undesirable phenomenon, since pumping of the ions in ion-gauges gave suspect information on the pressure elsewhere in the vacuum system. It was not until the 1950s that the possibility of utilizing the effect as a vacuum pump was exploited and the ion pump became a contender for ultrahigh vacuum applications.

The enhanced pumping stems from two basic factors; (1) the ionized molecules of the active gases such as O_2, N_2, H_2 and CO_2 have a much higher chemical activity than the neutral molecules and (2) the ions acquire considerable kinetic energy in the electric field and can therefore penetrate the surface on which they impinge. There are also some secondary effects as a result of sputtering, i.e. the removal of particles from the surface due to bombardment by the energetic ions. These sputtered particles can interact with the gas as they pass through it and/or can trap the gas at the surface on to which they deposit. Further, since the particles can be deposited over a wide area, the deposited layer can act as a getter surface similar to that of a sublimation pump.

The first prerequisite of an ion pump, therefore, is a method of ionizing the gas. At high pressures such as atmospheric pressure, the mean free path for electrons is very small and high fields are required to ionize the gas. At such pressures the main ion loss mechanism is recombination in the gas, so that the pumping action would be negligible. As the pressure is decreased the mean free path increases and the required field for ionization is reduced. At pressures around 1 Pa the ionization current is high and recombination takes place mainly at surfaces, effectively trapping the gas so that a good pumping action results. As the pressure is reduced further, say to 10^{-2} Pa, the mean free path of the electrons becomes large compared with the dimensions of the vacuum vessel (at 10^{-6} Pa the mean free path is of the order of 6 km) and the probability of an ionizing collision by an electron traversing the vessel is very

small. High fields are required and low currents result. Thus a simple discharge between two electrodes will have pumping action over a very limited pressure range around 1 Pa.

On the other hand, since the total amount of gas that can be trapped on the surface is limited, this type of pumping action is of more interest at lower pressures where 'saturation' effects do not present problems. Therefore, the design of ion pumps is aimed mainly at extending the pumping range down to lower pressures. To do this some means is required to provide adequate ionization at low pressures, preferably without excessively high voltages. The problem is identical to that of ion gauges and indeed there are similarities between the designs of ion pumps and of gauges.

There are two approaches which can be made to obtain sufficient ionization at low pressures: (1) the introduction of a large number of electrons into the system and (2) the electrons are made to traverse the system many times before finally being collected at the positive electrode. Since the efficiency of electron production is not very great, the electron current is usually restricted purely by power considerations. Therefore, although a good supply of primary electrons is often required, the main factor in the design of an ion pump is an arrangement for extending the electron paths.

There are several configurations of electrostatic fields, or combination of electrostatic fields with magnetic or r.f. fields, which will cause electrons to oscillate or spiral within a limited area. Many of these have been proposed as pumps or gauges but only a limited few have emerged as practical pumps for ultrahigh vacuum applications. These are described in this section.

The first use of ion pumping for ultrahigh vacuum was described by Alpert[41] when he used the hot cathode Bayard–Alpert ion gauge to pump small glass sealed-off systems, as part of the definitive studies of ultrahigh vacuum techniques at the Westinghouse Laboratories in the early 1950s. The Bayard–Alpert gauge (BAG) consists of a fine wire ion collector surrounded by an open mesh co-axial cylindrical grid, with the thermionic tungsten filaments mounted outside the grid. It is described in detail in Chapter 4 (Section 4.2.1). As a result of the electrostatic field, the electrons make several passes across the gauge before they are collected on the grid wires. *Figure 4.3* in Chapter 4 shows some typical trajectories. As a result there is an enhanced ionization efficiency. The ions formed outside the grid are mainly trapped on the envelope surface which is normally held at earth potential by a conducting film of tin oxide. The ion current to the envelope screen is about 5 to 10 times the collector currents in most gauges. Thus, by running the gauge at its highest current it acts as an ion pump. The average pump speed, however, is low, around 10^{-4} m^3s^{-1}, and the total capacity is limited to about 10^{14} molecules. On the other hand, since pumping continues down to 10^{-9} Pa, it is quite useful for small baked ultrahigh vacuum systems as a final pumping stage.

Apart from the use of the ion gauge, no practical ion pumps have emerged which rely on ion-pumping alone. Most pumps which are now available combine ion pumping with getter pumping, either by evaporating a metal getter film on to the pump walls (evapor-ion pumps) or by making use of the ions themselves to sputter a getter film (sputter-ion pumps).

The simplest evapor-ion pump is obtained by incorporating a means of evaporating a metal to form a getter film on the envelope of an ion gauge. A convenient method is to include a second filament in the gauge on which is

wound a titanium or zirconium wire, as in the sublimation pump. The pump can be used as a gauge until the zirconium coated filament is heated and the getter film deposited on the walls, thereafter the gauge readings may be affected. The getter film raises the pump speed to about 5×10^{-4} m^3s^{-1}. Such pumps were commercially available and were employed as throw-away appendage pumps for the final pump down of microwave tubes, etc. Once open to air they cannot normally be re-used, unless more than one evaporator filament is mounted in the same bulb. Kornelsen[42] describes an evapor-ion pump based on the inverted magnetron gauge[43] with a titanium evaporator in the form of wire wound around the tungsten anode. The pump speed for nitrogen was about 10^{-3} m^3s^{-1} but the significant factor was the improved capacity for inert gas pumping, by more than two orders of magnitude, as a result of evaporating the film of titanium on to the cathode which trapped the inert gas ions.

Larger designs of evapor-ion pumps based on the hot cathode ionization gauge, in the form of a triode with a suitable evaporator, have been described in the literature[44,45,46] with pump speeds up to 10 m^3s^{-1}. In many of the designs the titanium is continuously fed into the area from which it is evaporated and the metal envelope, normally water cooled, acts as the ion collector. Such pumps consume a large amount of power, are rather complex in structure and have a limited operational life before the titanium feed has to be replenished. It is not surprising therefore that commercial exploitation has been limited. An alternative approach to the triode evapor-ion pump is the orbitron pump. In this pump the long electron paths are achieved by launching the electrons into orbit in the electrostatic field set up between two concentric cylinders. The inner cylinder, in the form of a rod or wire small enough to avoid electron capture, is held positive and the outer cylinder negative. The actual applied potential depends on the diameters of the cylinders but is normally of the order of 5 kV.

Initial studies of the orbitron principle were carried out by Herb *et al.*[47] and it was first exploited as a vacuum gauge[48]. Some typical orbits calculated by Hooverman[49] are shown in *Figure 3.25*. These are calculated for a radial plane whereas in practice the electron will also have an axial velocity to give helix type orbits which can be reflected by end plates. The electron paths are considerably longer (more than 10^3 times those in a Bayard–Alpert gauge) and therefore give a much higher ion current for a similar electron current. The titanium is evaporated from a 'slug' mounted on the anode rod. The slug, having a larger diameter than the anode rod, intercepts the electron orbits and is heated by the bombarding electrons. The electrons are initiated from a thermionic cathode at one end of the cylinder.

A typical structure is shown diagrammatically in *Figure 3.26*. The pump speed is related to the ion current which in turn is related to the orbiting electron current. The electron current depends on the applied voltage and the geometry, in particular it is proportional to $L/\log(r_c/r_a)$ where L is the length of the pump and r_c and r_a are the radii of the cathode and anode respectively. Since r_a must be small to obtain long path length orbits, it is advantageous to have a long pump of small diameter. The resulting pumps are relatively compact, for example a 5×10^{-2} m^3s^{-1} pump could have an outside diameter of less than 10 cm with a length of about 25 cm. To achieve such a pumping speed a current of some 25 mA is required which implies a dissipation of

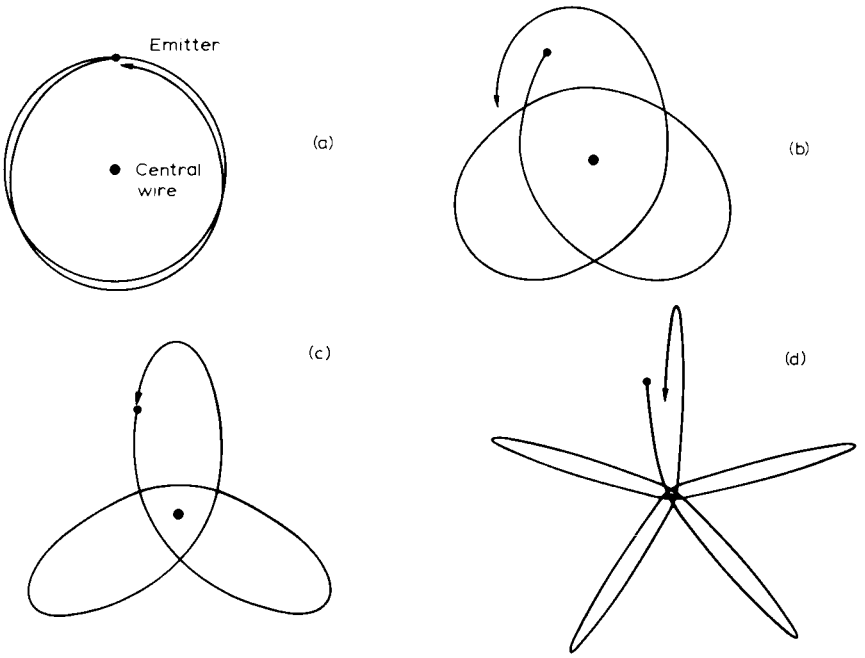


Figure 3.25 Example of electron orbits in an orbitron field according to Hooverman[49]

125 W, hence the need for water cooling the pump walls. Active gases are pumped by the titanium sublimation but the inert gases are pumped only by being trapped at the cathode as a result of ion burial. Since the field is logarithmic, the average energy of the ions on hitting the cathode will be less than a few hundred volts and consequently the pump speed for inert gases such as argon will be fairly low, typically 1% of the nitrogen pump speed.

An improvement is obtained with a triode configuration described by Bills[50] whereby the electrons are constrained to orbit between the anode and a concentric grid but the ions can pass through the grid and are accelerated to the cathode, which is negatively biased. This not only has the advantage of more efficient pumping of inert gases by burial in the cathode but also allows a larger gas conductance into the pump. Furthermore, it was found that the cathode did not have to be concentric with the other electrodes and that a multiple pump with several grid-anode structures in one envelope could be designed.

A pump construction with four cells is shown in *Figure 3.27*. This pump had a measured pump speed of 1.7 m^3s^{-1} for nitrogen and 2.5×10^{-2} m^3s^{-1} for argon when 4 kV was connected between the anode and grid and the cathode was 370 V negative with respect to the grid. In this pump the titanium was evaporated from a separate source which was heated continuously. The total power consumption was of the order of 1.5 kW.

Power dissipation seems to be one of the main disadvantages of the orbitron pump. It necessitates water cooling and also requires rather substantial power supplies. On the other hand the orbitron pump is compact, free from any magnetic fields, can pump down to 10^{-9} Pa at constant pump speed and it is

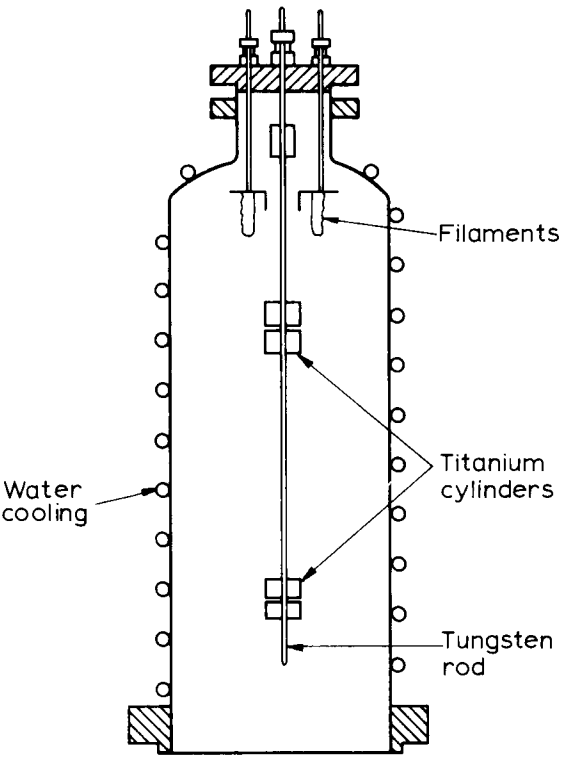


Figure 3.26 Sectional diagram of a typical orbitron pump

contamination free. It does, however, require a good backing pressure, $<10^{-1}$ Pa, to start it, and its pump speed decreases markedly at such pressures.

Although the orbitron pump became commercially available and found many applications, by far the most popular ion pump was and still is the sputter-ion pump. The basic design of the sputter-ion pump derives from the Penning cell, named after its inventor[51]. The cell consists of two parallel cathode plates with a cylindrical anode placed midway between them, having its axis normal to the plane of the plates (*Figure 3.28(a)*). A potential of a few kilovolts is applied between anode and cathodes and a magnetic field of the order of a kilogauss is applied in the direction of the anode axis. The structure is essentially an electron trap; electrons produced by ionization or by ion bombardment of the cathodes are forced to oscillate in the potential well between the cathodes. They are prevented from reaching the anode by the magnetic field which gives them long spiral paths and consequently high ionization probability. As a result a cold cathode discharge can be maintained down to pressures of the order of 10^{-9} Pa. The cell was originally exploited as a vacuum gauge[52] when the problem of pumping was observed. Its potential as a pump therefore was soon appreciated when Gurewitsch and Westendorp[53] showed that a greater pumping action could be obtained using chemically active cathodes such as titanium. However, it was not until 1958

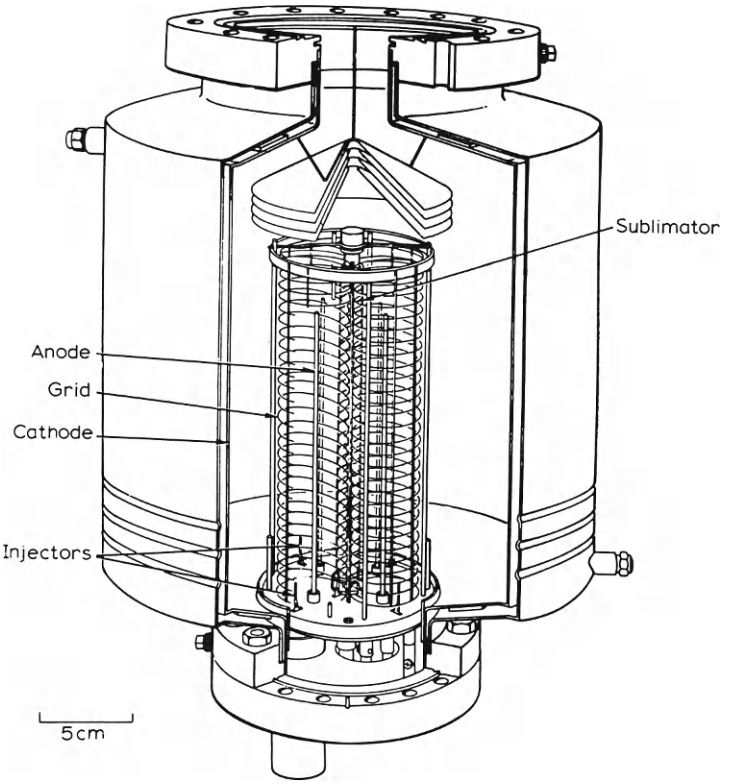


Figure 3.27 Diagram of orbitron type pump with grid electrodes according to Bills[50]

when Hall[54] conveniently combined several cells in parallel in one envelope to give high pumping speeds, that the Penning cell realized its potential as a practical ultrahigh vacuum pump. The pump design produced by Hall and shown in *Figure 3.28(b)* is the basic form for most sputter-ion pumps.

The pumping action is a result of several mechanisms which can be explained by reference to *Figure 3.29*. Ions produced in the discharge gain energies up to say 5 keV before striking the cathode and, as a result, sputter cathode material on to other surfaces of the pump. Because the ions are not produced uniformly across the cell and the field also is non-uniform, erosion caused by the sputtering is more pronounced in some areas than in others. The arrows of *Figure 3.29* represent material sputtering from the cathode, and the different shadings on the cathode and anode represent the areas where this material deposits.

Using titanium cathodes, chemically active gases may combine with the sputtered material in transit to the regions of deposition or become adsorbed after striking the deposits. The trapped gas is represented as black squares. Both active and inert gases accelerated to the cathode as ions may become implanted there, represented by the black and white triangles. However, continued sputtering will again release the uncombined gas, so that only on those regions of the cathode where there are growing deposits can inert gas be

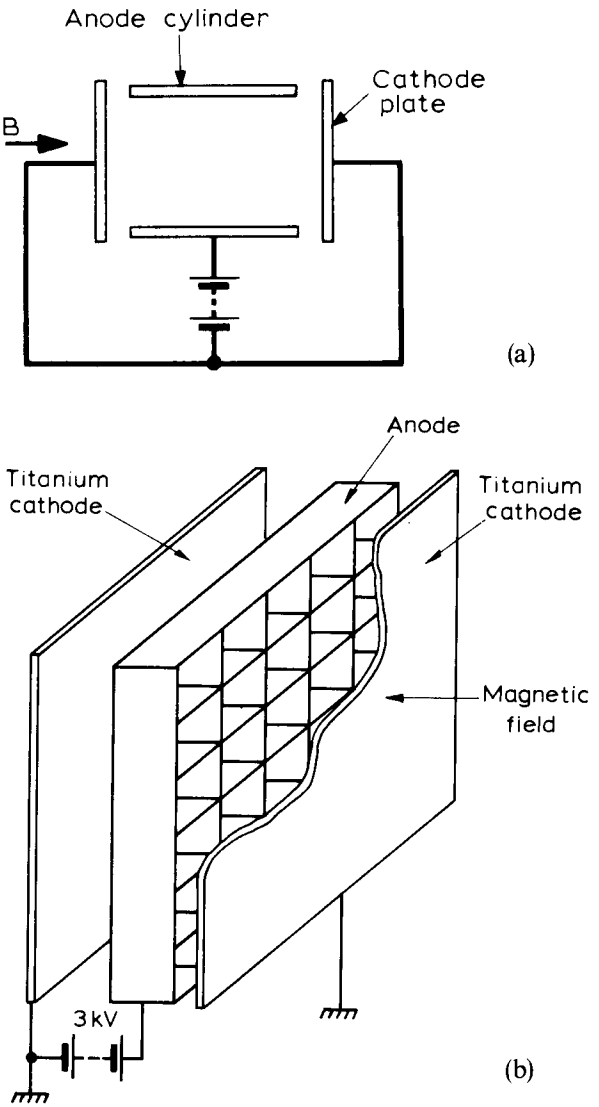


Figure 3.28 (a) The basic Penning cell; (b) Sputter ion pump basic construction, Hall[54]

permanently trapped. Auto-radiography of cathodes[55], from pumps which have pumped radioactive krypton confirm this.

Autoradiography also has shown that some inert gas is absorbed on the anode. A mechanism which can explain this has been proposed by Jepsen[56], whereby the ions striking the cathode are reflected under suitable conditions as energetic neutrals which could then impinge on the anode and become buried. The energy of the neutrals depends on the angle of incidence and the atomic weight of the cathode material. For titanium bombarded at normal incidence the effective energy of the reflected inert gas atoms is small. The

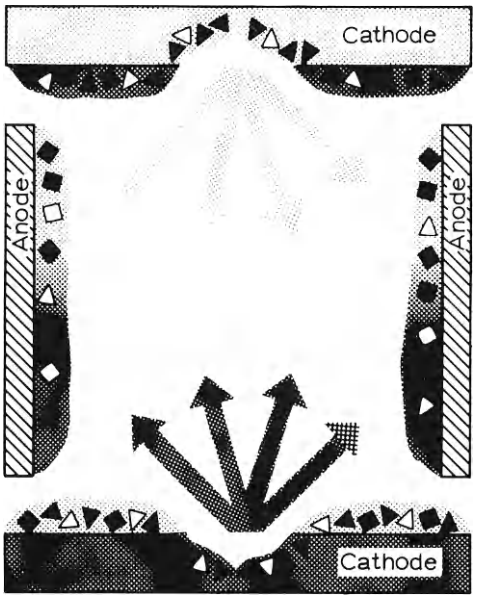


Figure 3.29 Schematic diagram of the pumping mechanism of a sputter-ion pump

pump speed for inert gases is therefore low, typically for argon it is 1% of that for nitrogen in the Hall design of pump. The pumping of hydrogen is also interesting. Hydrogen is readily adsorbed by titanium, forming pseudo hydrides. However, the hydrides dissociate as the temperature is raised and the sputter yield for hydrogen is very low. When pumping hydrogen the pump speed will therefore drop with time, as equilibrium is reached between absorption and hydride dissociation. There is a further problem in the sputter-ion pump with inert gases, namely the release of gas from the cathode. This leads to fluctuations in pressure when pumping atmospheres containing inert gases, particularly argon. Known as 'argon cycles' these fluctuations can result in periodic gas bursts, with pressure rises of more than an order of magnitude.

The low pump speed and fluctuation with argon has been considered so important that much of the pump development since the initial design of Hall has been directed towards improving these characteristics.

Since the major pumping action for inert gases was considered to be due to ion burial in the cathode under sputter deposits, the designs were aimed at improving this facility. Brubaker[57] suggested a triode design which aimed to enhance sputtering by causing ions to strike the cathode at oblique incidence and also provided a collector surface at a potential midway between cathode and anode potential where lower energy ions could be buried without risk of

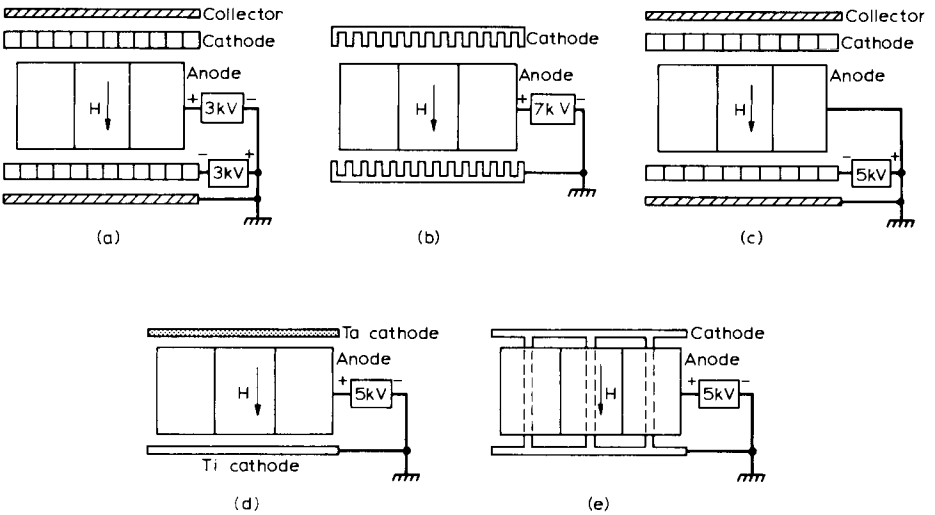


Figure 3.30 Pump designs for improved inert gas pumping. (a) Triode, Brubaker[57]; (b) Slotted cathode diode, Jepson *et al.*[59]; (c) Triode, Hamilton[58]; (d) Differential diode, Tom and James[60]; (e) Magnetron, Andrew *et al.*[55]

further sputtering. The triode design is shown diagrammatically in *Figure 3.30(a)*. In later designs it was found that the triode construction functioned equally well with the collector connected to anode potential[58], *Figure 3.30(c)*, which is of course in conflict with the suggested pumping mechanism of enhanced ion burial. However, at oblique incidence the scattered neutrals in the forward direction will have a significant fraction of the energy of the incident ion. So it seems reasonable to assume that in the triode pump the inert gas is pumped as energetic neutrals rather than as ions. Argon pump speeds up to 25% of those for nitrogen are claimed, with absence of argon cycles even when pumping pure argon. However, because of the greater depth in the direction of the magnetic field, a lower total pump speed is obtained than for a similar sized diode pump or alternatively a stronger magnet is required. Also the complexity of the structure was criticized, and an alternative slotted cathode design was proposed by Jepsen *et al.*[59] (see *Figure 3.30(b)*). The enhanced argon pumping over a plain cathode was explained by a greater yield of sputtered material from the oblique bombardment of the sides of the slots coupled with greater area of cathode; of particular importance is the base of the slots, where deposition and trapping occurred. Although the pump speed for argon was less than for the triode, 6–10% of the nitrogen speed, and argon cycles could occur when pumping pure argon, the overall pump speed was good and the structure simple to make and to operate.

A further development occurred when alternative materials were used for the cathodes, particularly when a different material was used for each cathode, the so-called differential pump of Tom[60] illustrated in *Figure 3.30(d)*. The pump uses one cathode of titanium and one of tantalum, and the improved inert gas pumping was originally ascribed to the different sputter rates of the two cathodes resulting in a build-up on one of them. This explanation, however, was not satisfactory, since it is clear that such a build-up of the deposited layer would eventually result in the cathodes having surfaces of

similar material, and in any case the sputter rates of titanium and tantalum are very similar. Again the trapping of energetic neutrals seems a more likely explanation. The energy of the scattered neutrals depends on the atomic weight of the metal and tantalum, being a much heavier metal than titanium, would give scattered neutrals with significant energy even from normal incidence. The titanium is presumably still required to getter the active gases. An alternative structure, whereby a pellet of tantalum was attached to the titanium cathode plates at the centre of each cell, has been described by Bächler[61,62] and is known as the pill cathode.

A design based on the magnetron cell was also found to give enhanced pumping of argon. The magnetron cell is similar to the Penning cell but with a rod mounted axially down the anode cylinder linking the two cathodes, see *Figure 3.30(e)*. The pump performance and the pumping mechanism have been described by Andrew *et al.*[55]. The effect of the rod on the field is such that 90% of the ion current goes to the rod and, since the ions impinge at oblique angles, the sputter rate is high. A build-up of sputtered deposit takes place virtually over the entire surface of the cathode plates, trapping the incident inert-gas ions. This was confirmed by pumping radioactive krypton where 84% of the gas was found to be absorbed on the cathode plates. The argon pump speed was approximately 15% of that for nitrogen. A modified construction, in which the rod did not extend right across the cell but was in the form of separate 'posts' protruding from each cathode, gave a similar performance[63]. A further enhancement of argon pumping was obtained by using tantalum posts, which facilitates the trapping of energetic neutrals. A similar effect was found in the magnetron pump[64].

The magnetron design was initially investigated in order to take advantage of the known ability of magnetron cells to support discharges at pressures below 10^{-10} Pa; for a standard diode the current and pump speed dropped appreciably at pressures below 10^{-8} Pa. Another innovation has been to use one cathode of high vapour pressure metal in the differential pump so that the metal atoms evaporated and/or sputtered contribute to the ionization current[65]. Using a magnesium cathode a 50% increase in pumping speed is claimed over the whole operational pressure range. Because of the high vapour pressure, however, the pump should not be baked above 350°C. Whatever type of cell is used, diode, triode or magnetron, there are certain design factors which must be considered. For example there is an optimum cell size both in diameter and length. Also the gap between the anode cylinder and cathode plates is a compromise between adequate conductance and provision of magnetic flux. Andrew[64] has given data showing how the pump speed of an array of cells depends on the anode cathode conductance. In particular he showed that as the distance of a cell from the pump orifice increased, so the pump speed of the cell decreased and indeed the cells furthest away become ineffective unless the conductance is large. Thus there is a limitation to the extent that the anode array of *Figure 3.28(b)* can be increased to obtain higher pump speeds. For the larger pumps a construction of pump modules grouped round a central duct is normally employed, see *Figure 3.31*.

Provision of the magnetic field also requires some consideration. Since it is advantageous to bake the pump to say 400°C for outgassing and one does not always wish to remove the magnets whilst baking, the magnetic material is restricted to alloys such as Alnico or Ticonal or to ferrite magnets such as

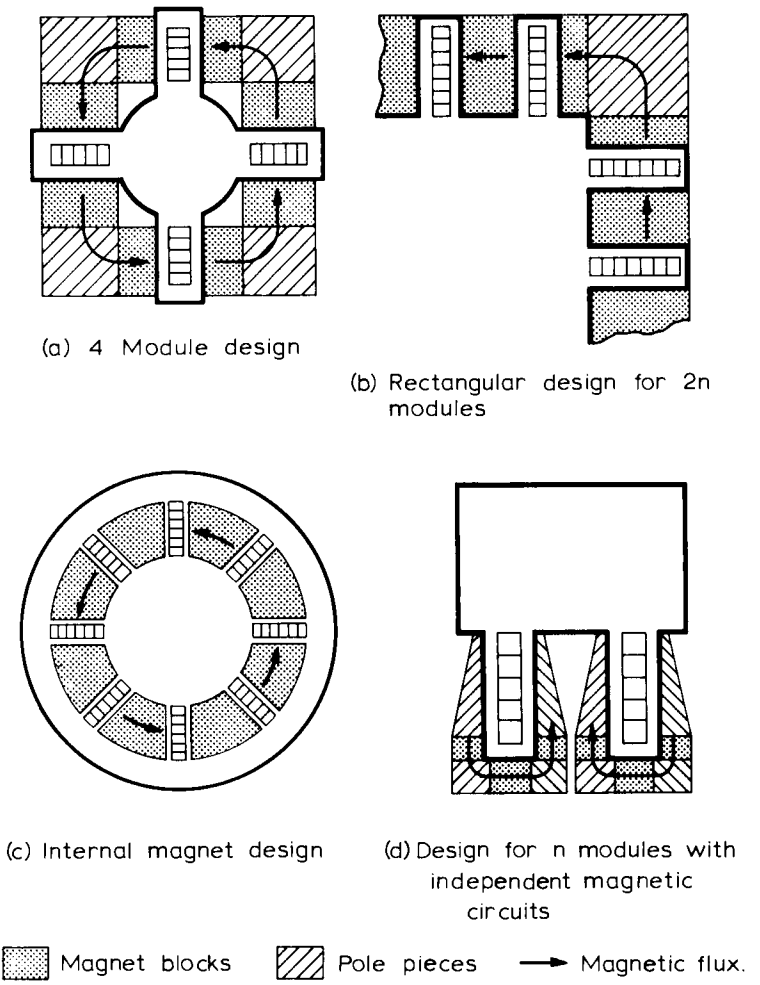


Figure 3.31 Module designs for large sputter ion pumps. (a) Four module design; (b) Design for 2*n* modules; (c) Internal magnet design; (d) Design for modules with independent magnetic circuits

Magnadur 3. The high coercivity of ferrite magnets makes them less sensitive to demagnetizing forces if mishandled. Although the flux density falls off markedly with temperature it is fully recoverable at room temperature for excursions up to 400°C. The alloy magnets whilst having a smaller temperature coefficient, will lose some of their flux if taken above 200°C. The ferrites are particularly suitable where a closed loop of magnetic flux can be obtained, *Figure 3.31(a)*. Such an arrangement gives the maximum field with the minimum amount of magnetic material, and also minimizes the stray magnetic field. For single or double modules, *Figure 3.31(d)*, alloy magnets seem to be preferred.

The power supply is also important. At low pressures, $< 10^{-3}$ Pa, a voltage of 2–3 kV is required to maintain the discharge and the pump speed increases with the field, see *Figure 3.32*. At higher pressures the pump will operate at

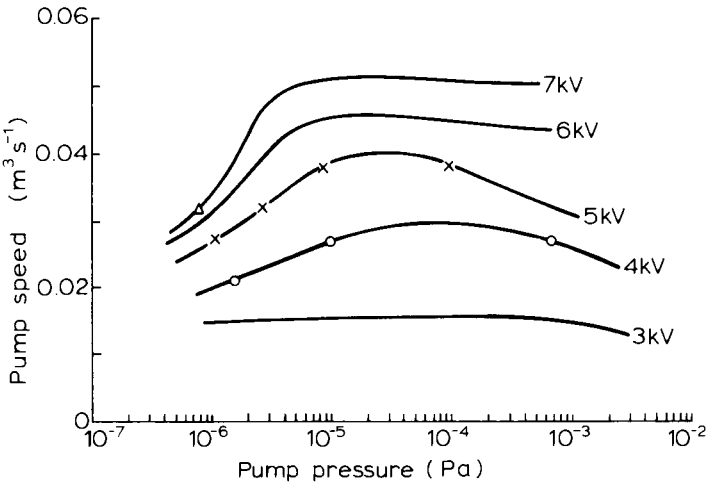


Figure 3.32 Pump speed as a function of pressure for a magnetron pump showing the effect of the HT voltage

much lower voltages, but with high currents of the order of 10^{-1} A. A compromise is required between optimum voltage and current conditions over the entire pressure range and the cost of the supply. The most economical method of power generation that is capable of achieving characteristics close to optimum is by means of a bridge rectification circuit fed from a transformer with a high leakage reactance. Similarly the maximum voltage is also chosen from the point of view of economics not only in the generation of the voltage but also in the insulation of the pump and connections. Normally potentials from 3–5 kV are employed.

Because several mechanisms are involved in the pumping process, the pump speed will reach an equilibrium value only when the getter film is saturated and there is a balance between the gas released from the cathode by erosion and the gas implanted as ions or neutrals. The latter process depends on the nature of the gas and the quantity already implanted in the cathode. It can, therefore, be expected that the speed of the pump will show saturation effects and will be influenced by the past history of the pump. Measurements of pump speed have shown that the saturation or equilibrium pump speed is reasonably reproducible[66], varying with pressure as shown in the typical curves of *Figure 3.32*. Before equilibrium is reached, however, much higher pump speeds can be attained and the experienced operator can utilize this effect to advantage. The pump does not saturate until at least 10^{-1} Pa m^3 have been pumped but at 10^{-7} Pa this could mean thousands of hours of pumping. Activation of the pump to a non-equilibrium condition can be achieved by what is normally termed a regenerative bake, that is, baking the pump to around 250°C whilst it is operating. The high temperature is thought to increase the sputtering yield and allow diffusion of the trapped gas from the surface. An alternative method is to run the pump with an argon atmosphere of 10^{-3} Pa which produces a clean titanium sputtered layer. However, the increase in argon residual thereafter is not always acceptable.

The sputter-ion pump can be started at pressures around 1 Pa, but at such

pressures the pump speed is very low and the dissipation high. As a result the pump will heat up and could degas at a higher rate than it pumps. It is therefore better to rough out to at least 10^{-1} Pa before operating the ion pump. Also at the high pressures the cathode life is limited due to the greatly increased sputter rate. At pressures of 10^{-4} Pa lives of the order of 50 000 h are usually cited.

The main advantage of ion-pumps is their freedom from contamination such as hydrocarbons. This means that they do not have to be provided with cold traps or baffles and all the available speed of the pump can be utilized, to the extent of mounting the pump within the working chamber. They are capable of attaining pressures below 10^{-8} Pa and are simple to use. For most pumps, a source of electricity is all that is required and they will operate unattended without the need for safety devices. The ion current of the pump is a useful indication of pressure and with slight modification to the circuits it is possible to use the ion pump as its own leak detector. To exploit the clean pumping feature of an ion pump it is best backed by sorption pumps. It can, however, be used with rotary pumps provided they are suitably trapped. Once the backing pump has taken the pressure down to 10^{-1} Pa it can then be isolated from the system. The disadvantage of the ion-pump is the limited capacity and the low pump speeds for inert gases. For the sputter ion pump the strong magnetic field can be an embarrassment in some applications.

3.8 Choice of pumps

As a result of the different pump types that have been developed for ultrahigh vacuum applications, the vacuum engineer is faced not only with a choice of similar pumps from the different manufacturers but also a choice of pumping techniques. In order to assess the advantages and disadvantages of any pump or combination of pumps a reliable criterion of performance is required. Pump speed is the obvious choice but it is important that like is compared with like. Pump speed can be a variable quantity depending on the gas pressure and species, the history of the pump and, more importantly, the way it is measured. Quite large errors can occur in the measured value of pump speed if gas flow effects prevent homogeneous distribution of the molecules. Because of this, standard conditions for such measurements have been recommended internationally and are aimed at ensuring reliable and reproducible values. Normally, to overcome errors in gas flow and pressure measurements resulting from gas streaming effects, large vessels are required. However, by judicial positioning of the gauges and orifices, etc., large chamber conditions can be simulated with a much smaller test dome. *Figure 3.33(a)* illustrates the recommended test dome which was first introduced for diffusion pumps. The important criteria are the length to radius ratio and the direction of the gas inlet pipe. The flow rate through the inlet pipe is measured whilst maintaining the pressure at the pump (i.e. in the dome) constant. Later the recommendations were extended to measurements, at much lower pressures, where the test dome and gauges could be baked out. *Figure 3.33(b)* shows the suggested test dome for such measurements, it consists of two chambers separated by an orifice plate which facilitates measurement of the gas flow. Assuming molecular flow conditions the pump speed is given by

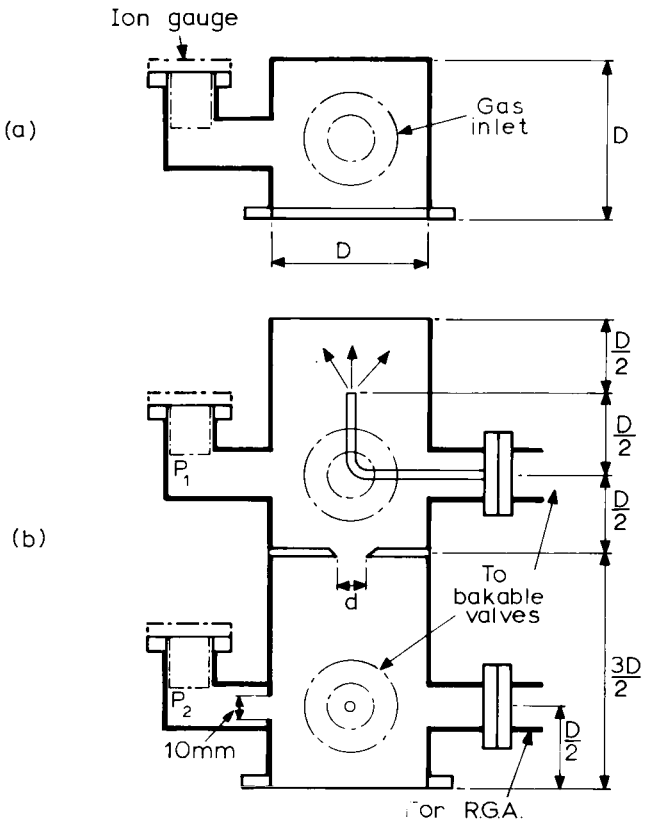


Figure 3.33 Test dome for pump speed measurement. (a) For diffusion pumps down to 10^{-5} Pa; (b) For ultrahigh vacuum pressures

$$S = C\{\alpha(P_1/P_2) - 1\} \tag{3.9}$$

where α is a correction factor to take into account errors in the two gauge sensitivities and is measured in a separate experiment, P_1 is the measured pressure in the upper chamber, P_2 that in the lower chamber and C is the conductivity of the orifice which can be calculated from the geometry (see Section 1.3.3). Typically the diameter, d, of the orifice should be in the range of 0.05 to 0.1 D. For more details see Steckelmacher[67].

Most manufacturers adhere to the recommendations in principle and although Feng and Xu[68] suggest that modification of the dome dimensions would give a more accurate measurement of the intrinsic pumping speed, the values of pump speed and ultimate pressure quoted in commercial data sheets can usually be taken as accurate enough for comparison purposes. When comparing different types of pumps, however, the characteristics of the pump in terms of variation in speed with gas pressure and gas species, etc., discussed in the previous sections, must be taken into account. The choice is not always an easy one. For applications where a large amount of inert gas has to be handled, as for example in sputtering chambers, it is clear that the pumps retaining the gas within the pump body and especially those depending on

chemical adsorption such as ion pumps and getter pumps are not satisfactory. For many applications the choice is less clear and it is often one of personal preference or prejudice. Factors such as supply requirements, size and cost often outweigh the advantage in pump performance that one system may have over another.

In *Table 3.5* an attempt has been made to list and compare parameters of interest in the assessment of pumps. Because there are differences in construction, performance and cost of similar pumps from different manufacturers, the data should only be considered as a rough guide. Nevertheless it gives an overall picture of the advantages and disadvantages of various types, which acts as a useful starting point when considering the most suitable pumping system for a particular application.

Although pump speed is a reasonable criterion for most pumps, if one is considering sorption pumps or sublimation pumps this is not very meaningful. In order to make a fairer comparison the sort of pump required for a large vacuum system of around 0.5 m^3 volume is considered in *Table 3.5*. For some types an optimum size pump is not available and a pump larger than required is listed. However, this should not affect the general picture.

For any ultrahigh vacuum system a combination of pumps is required to attain the low pressure. For systems where the gas throughput is likely to be high, for example in systems where sputtering or ion etching forms part of the processing, then the diffusion pump backed by a suitable rotary pump offers probably the most economic solution, provided adequate trapping is employed. However, one must not lose sight of the fact that an efficient baffle and trap for the oil vapour also reduces the pump speed by at least 50%. Also, even with the best trapping arrangement, there is always the slight danger of hydrocarbons contaminating the system. The alternative pumping system for this application is the turbomolecular pump backed by a rotary pump and it is becoming more popular with the improved pumps now available. It is more expensive and, although reliable, to some extent it is less robust. Certainly an accidental ingress of a particle into the pump could be costly. For ultimate pressures $<10^{-8}$ Pa the turbomolecular pump requires a lower backing pressure than conveniently given by a standard rotary pump and often requires a diffusion pump to back it. However, an alternative arrangement is to use the turbomolecular pump in conjunction with an ion pump. The turbomolecular pump handles the bulk of the gas and the ion pump is turned on to achieve a better ultimate pressure (see Chapter 7).

In systems where large gas throughput is not required the ion pump comes into its own. Although more expensive than a diffusion pump, it does not require trapping and full advantage can be taken of its pumping speed. It is completely hydrocarbon free and to exploit this, it is best backed with sorption pumps. One of the objections to the ion pump system is the large percentage of hydrogen in the residual gas which could be unacceptable for certain applications and the slow pumping speed for inert gases. There are also applications where the stray magnetic field from the sputter-ion pump cannot be tolerated. Also it must be pointed out that the sorption pumps have a limited capacity; for a single pump approximately 100 g of zeolite is required for each 10^{-3} m^3 (litre) volume to be pumped. For the example given in *Table 3.5*, 50 kg of zeolite would be required. Clearly this would not be an acceptable solution, and for large volumes, cycles of sorption, and desorption at room

TABLE 3.5. Comparison of pump parameters for a vacuum system of about 0.5 m^3 volume

Pump	*Pump speed* ($m^3 s^{-1}$) N_2	H_2	A	*Backing pressure* (Pa)	*Weight* (kg)	*Size* ($m^3 \times 10^{-2}$)	*Total capacity* (Pa m^3)	*Services* W = water E = electricity LN = liquid N_2	*1983 cost* (£)	*Remarks*
Diffusion	0.7	1.3	0.7	1.0	20	2.5	Unlimited	E, W	1200	Water cooled baffle included
Turbomolecular (drive unit)	0.5	0.5		10^{-1}	30 (3)	2.5 (0.3)	Unlimited	E, W	4000 (650)	
Sputter ion (power supply)	0.5	1.0	0.05	10^{-1}	150 (40)	8.0 (3.0)	10^4	E	3000 (1500)	Ti readily absorbs H_2 but may release it later
Cryogenic (refrig.) (compressor)	0.7	0.5	0.5	10^{-1}	19 (65)	1.0 (10)	10^3	E	7000	Can be generated by baking
Sublimation (power supply)	~1.0	~1.0	Nil	10^{-2}	3 (17)	1.8 (1.5)	10^2	E, W	300 (600)	Does not include water-cooled cylinder
Rotary	0.003	0.003	0.003	10^5	90	4.0	Unlimited	E	450	Backing pump
Sorption (2 in series)	2000 g Zeolite			10^5	30	1.5	10^4	LN	500	The pumps would need to be used sequentially

temperature would be used, i.e. pumps would be used alternatively, whilst one, pumping at liquid nitrogen temperature, is open to the system, a second, exposed to the atmosphere, would be warmed up.

For large systems, such as space simulation chambers, which cannot be easily baked, the shear bulk of the necessary pumps is a problem. Cryogenic pumps offer a solution here, especially if they are built into the system. However, they still require the vacuum system to be pumped out to a pressure below 10^{-1} Pa before they can be efficiently operated, which often requires a more sophisticated system than a simple rotary pump. Sublimation pumps are also very useful in this application but of course they will not pump inert gas. The self-contained cryogenerator now offers a compact pump for smaller systems with desirable attributes such as cleanliness and high speed for all gases, but the capital outlay is at present high. It can be used with sorption pumps in a similar procedure to the ion-pump, or in conjunction with a trapped two-stage rotary pump. Since the roughing-out time is short the possible hydrocarbon contamination is minimal.

In conclusion, the ideal pump evacuating from atmosphere to 10^{-8} Pa is not yet available but provided the influx of gas into the system from desorption, etc. is kept to a minimum, modern, commercially available vacuum pumps used in a suitable combination allow ultrahigh vacuum pressures to be achieved readily and consistently.

3.9 References

1. GAEDE, W., *Ann. Physik*, **41**, 337, (1913)
2. PIRANI, M. and YARWOOD, J., *Principles of Vacuum Engineering*, Chapman and Hall, (1961)
3. HABLANIAN, M. H. and MALIAKAL, J. C., *J. Vac. Sci. Technol.*, **10**, 58, (1973)
4. POWER, B. D. and CRAWLEY, D. J., *Vacuum*, **4**, 415, (1954). (Published 1957)
5. RIDDIFORD, L., *Vacuum*, **3**, 49, (1953). (Published 1954)
6. CRAWLEY, D. J., TOLMIE, E. D. and HUNTRESS, A. R., *Trans. of 9th Nat. Vac. Symposium 1962*, 399, Macmillan, (1962)
7. VENEMA, A. and BANDRINGA, M., *Philips Tech. Rev.*, **20**, 145, (1959)
8. LAURENSON, L., *Vacuum*, **30**, 275, (1980). [See also O'HANLON, J. F., *J. Vac. Sci. Technol.*, **A2**, 174, (1984)]
9. HOLLAND, L., *Vacuum*, **20**, 175, (1970)
10. BAKER, M. and LAURENSON, L., *Vacuum*, **16**, 633, (1966)
11. HOLWECK, F., *Comp. Rend.*, **177**, 43, (1923)
12. VON FRIESEN, S., *Rev. Sci. Instrum.*, **11**, 362, (1940)
13. BECKER, W., *Vakuum Tech.*, **7**, 149, (1958)
14. KRUGER, C. H. and SHAPIRO, A. H., *Trans. 7th Nat. Vac. Symposium 1960*, 6, Pergamon, (1961)
15. OSTERSTROM, G. E. and SHAPIRO, A. H., *J. Vac. Sci. Technol.*, **9**, 405, (1972)
16. MIRGEL, K. H., *J. Vac. Sci. Technol.*, **9**, 408, (1972)
17. MAURICE, L., *Proc. 6th Internat. Vac. Cong. 1974, Japan, J. App. Phys. Suppl. 2 Part 1*, 21, (1974)
18. DUSHMAN, S., *Scientific Foundations of Vacuum Technique*, Wiley, (1962)
19. TURNER, F. T. and FEINLEIB, M., *Trans. 8th Nat. Vac. Symposium and 2nd Internat. Vac. Congress 1961*, **1**, 300, Pergamon, (1962)
20. READ, P. L., *Vacuum*, **13**, 271, (1963)
21. CREEK, D. M., PETTY, R. and JONATHAN, N., *J. Sci. Instrum.*, (*J. Phys. E.*) 2nd Series, **1**, 582, (1968)
22. HALAMA, H. J. and AGGUS, J. R., *J. Vac. Sci. Technol.*, **11**, 333, (1974)
23. HAEFER, R. A., *Le Vide*, **25**, 65, (1970)

24. BENVENUTI, C., *J. Vac. Sci. Technol.*, **11**, 591, (1974)
25. COUPLAND, J. R. and HAMMOND, D. P., *Vacuum*, **32**, 613, (1982)
26. HENGEVOSS, J., REISINGER, H. and WÖSSNER, H., *J. Vac. Sci. Technol.*, **7**, 251, (1970)
27. KLIPPING, G. and MASCHER, W., *Vakuum Tech.*, **11**, 81, (1962)
28. MOORE, R. W., *Trans. 8th Nat. Vac. Symposium and 2nd Internat. Vac. Congress, 1961*, **1**, 426, Pergamon, (1962)
29. HANDS, B. A., *Vacuum*, **32**, 603, (1982)
30. DAVID, R. and VENEMA, A., *Proc. 3rd Internat. Vac. Congress 1965*, **II**, 577, Pergamon, (1966)
31. GIFFORD, W. E. and McMAHON, H. O., *Proc. 10th Internat. Cong. Refrig.*, 1, (1959)
32. BENTLEY, P. D., *Vacuum*, **30**, 145, (1980)
33. CLOUD, R. W., BECKMAN, L. and TRUMP, J. G., *Rev. Sci. Instrum.*, **28**, 889, (1957)
34. CLAUSING, R. E., *Trans. 8th Nat. Vac. Symposium and 2nd Internat. Vac. Congress 1961*, **1**, 345, Pergamon, (1962)
35. McCRACKEN, G. M. and PASHLEY, N. A., *J. Vac. Sci. Technol.*, **3**, 96, (1966)
36. LAWSON, R. W. and WOODWARD, J. N., *Vacuum*, **17**, 205, (1967)
37. DELLA PORTA, P., GIORGI, T., ORIGLIO, S. and RICCA, F., *Trans. 8th Nat. Vac. Symposium and 2nd Internat. Vac. Congress 1961*, **1**, 229, Pergamon, (1962)
38. BAROSI, A., *Residual Gases in Electron Tubes* (Proc. 4th Internat. Conf.), Academic Press, London, 221, (1972)
39. DELLA PORTA, P. and FERRARIO, B., *Proc. 4th Internat. Vac. Congress 1968*, **1**, 369, Inst. of Physics, (1968)
40. TUCK, R. A., *Vacuum*, **22**, 409, (1972)
41. ALPERT, D., *J. Appl. Phys.*, **24**, 860, (1953)
42. KORNELSEN, E. V., *Trans. 7th Nat. Vac. Symposium 1960*, 29, Pergamon, (1961)
43. HOBSON, J. P. and REDHEAD, P. A., *Can. J. Phys.*, **36**, 271, (1958).
44. HERB, R. G., *Advances in Vac. Sci. & Technol., Proc. 1st Internat. Vac. Congress 1958*, **1**, 45, Pergamon, (1960)
45. HUBER, H. and WARNECKE, M., *Le Vide*, **13**, 84, (1958)
46. HOLLAND, L., LAURENSON, L. and HOLDEN, J. T., *Nature*, **182**, 851, (1958)
47. HERB, R. G., PAULY, T. and FISHER, K. J., (Abstract), *Bull. Am. Phys. Soc.*, **8**, 336, (1963)
48. MAURAD, W. G., PAULY, T. and HERB, R. G., *Rev. Sci. Instrum.*, **35**, 661, (1964)
49. HOOVERMAN, R. H., *J. Appl. Phys.*, **34**, 3505, (1963)
50. BILLS, D. G., *J. Vac. Sci. Technol.*, **4**, 149, (1967)
51. PENNING, F. M., *Philips Tech. Rev.*, **2**, 201, (1937)
52. PENNING, F. M. and NIENHUIS, K., *Philips Tech. Rev.*, **11**, 116, (1949)
53. GUREWITSCH, A. M. and WESTENDORP, W. F., *Rev. Sci. Instrum.*, **25**, 389, (1954)
54. HALL, L. D., *Rev. Sci. Instrum.*, **29**, 367, (1958)
55. ANDREW, D., SETHNA, D. R. and WESTON, G. F., *Proc. 4th Internat. Vac. Congress 1968*, **1**, 337, Inst. of Physics, (1968)
56. JEPSEN, R. L., *Proc. 4th Internat. Vac. Congress 1968*, **1**, 317, Inst. of Physics, (1968)
57. BRUBAKER, W. M., *Trans. 6th Nat. Vac. Symposium 1959*, 302, Pergamon, (1960)
58. HAMILTON, A. R., *Trans. 8th Nat. Vac. Symposium and 2nd Internat. Vac. Congress 1961*, **1**, 388, Pergamon, (1962)
59. JEPSEN, R. L., FRANCIS, A. B., RUTHERFORD, S. L. and KIETZMANN, B. E., *Trans. 7th Nat. Vac. Symposium 1960*, 45, Pergamon, (1961)
60. TOM, T. and JAMES, B. D., *J. Vac. Sci. Technol.*, **6**, 304, (1969)
61. BÄCHLER, W., *Vakuum Tech.*, **17**, 59, (1968)
62. BÄCHLER, W. and HENNING, H., *Proc. 4th Internat. Vac. Congress 1968*, **1**, 365, Inst. of Physics, (1968)
63. LAMONT, L. T., *J. Vac. Sci. Technol.*, **6**, 47, (1969)
64. ANDREW, D., *Brit. J. Appl. Phys.*, (*J. Phys. D.*) 2nd Series, **2**, 1609, (1969)
65. TOM, T., *J. Vac. Sci. Technol.*, **9**, 383, (1972)
66. ANDREW, D., *Vacuum*, **16**, 653, (1966)
67. STECKELMACHER, W., *Proc. 4th Internat. Vac. Congress 1968*, **1**, 67, Inst. of Physics, (1968)
68. FENG, Y. G. and XU, T. W., *Vacuum*, **30**, 377, (1980)

Chapter 4

Total pressure measurements

4.1 Introduction

The extensive studies during the 1950s to attain better vacuum conditions were dependent on the ability to measure and analyse residual gases at pressures below 10^{-5} Pa. It had long been considered that diffusion pumps combined with gettering could pump down to below such pressures but without the gauges to measure the pressure, the premise was little more than conjecture.

Gauges then available which depended on mechanical forces such as the mercury manometer and diaphragm gauges, or on thermal conductivity such as the Pirani gauge, become inoperative below 10^{-3} Pa due to the sparseness of gas molecules. Methods of compressing the gas before measurement to raise the apparent pressure, for example in the McLeod gauge, extended the range but only by an order or two. Some gauges depending on momentum transfer had extended the range down further but they were delicate and limited in application. One was then left with ionization gauges. Although they appeared capable theoretically of measuring much lower pressures, it was recognized as early as 1937 that the ionization gauges available also had a low pressure limitation, of around 10^{-5} Pa.

It was not until the work of Bayard and Alpert[1] in 1950, that an ionization gauge was devised whereby this limitation was significantly reduced and pressures below 10^{-5} Pa could be measured. The Bayard–Alpert gauge (BAG) represented a breakthrough in low-pressure measurement and gave the impetus for the vacuum research and development of the 1950–60s. Improvements in the BAG and other designs of ionization gauge followed, which made possible pressure measurements down to 10^{-10} Pa. Some of these were developed into commercial designs and are available today. Since these gauges cover the ultrahigh vacuum region down to 10^{-10} Pa there is no great need for further advancement and most of the work on gauges since 1970 has been aimed at consolidation, taking a closer look at existing designs and making refinements to improve on performance, reliability, packaging or cost.

The conventional triode ionization gauge in use prior to the 1950s was limited by residual currents at low pressures, mainly ascribed to soft X-rays produced at the grid by the incident electrons, which in turn produced a

photocurrent at the collector[2]. The improved gauges were aimed at reducing the residual current and/or increasing the sensitivity. Various configurations have been proposed for ultrahigh vacuum gauges but the BAG remains today the most used total pressure gauge for high and ultrahigh vacuum applications. However, although the ionization gauge is universally employed for the measurement of pressures in the ultrahigh vacuum range, it is by no means ideal. It is not an absolute gauge and its sensitivity is dependent on the composition of the residual gas being measured. It must therefore be calibrated for a known gas and its relative sensitivity for other gases must be determined. Also, the gauge reacts with the gas environment to the extent that much of the investigation of ion gauge performance is devoted to studying physical and chemical processes in the gauge.

An absolute gauge with a pressure range extending into the ultrahigh vacuum region, which at the same time does not affect the environment, would obviously be of great interest. One aspect of the vacuum environment which offers this possibility is the momentum transfer that can take place between two surfaces via the gas molecules. The momentum may be transferred from a moving surface to a stationary surface or between surfaces of different temperatures, the so-called radiometer effects. Viscosity gauges exploiting the former effect and Knudsen type gauges exploiting the latter, have been constructed which are capable of measuring pressures in the ultrahigh vacuum region. In both cases the forces to be measured are very small, at least an order down on the forces due to the gas pressure alone, i.e. less than 10^{-6} N m^{-2}. A gauge sensitive to such small forces is bound to be rather fragile and although a number of designs for practical gauges have been described, in general they have not been robust or simple enough for general use. However, because they are virtually independent of the gas composition and do not interact with the gas in any way, they present useful gauges as 'standards' for calibration purposes.

Apart from describing the mechanism and performance of the various ultrahigh vacuum gauges, this chapter also deals with the problem of gauge calibration at such low pressures and the accuracy that can be expected.

4.2 Ionization gauges

If a gas is ionized by electron impact from an electron current passing through it, then the number of positive ions formed is directly proportional to the molecular density ρ

$$i_+ = C\rho i_- \tag{4.1}$$

where i_+ is the ion current, i_- the electron current and C the constant of proportionality. When there is equilibrium between pressure and density at temperature T, then the well known gas–kinetic equation $P = \rho kT$, in which k is Boltzmann's constant, is valid. Equation (4.1) can then be written as

$$P = \frac{1}{K}\frac{i_+}{i_-} \tag{4.2}$$

where $K = C/kT$ and is known as the gauge sensitivity. Thus, to measure pressure by ionization requires a source of electrons (cathode), an accelerating

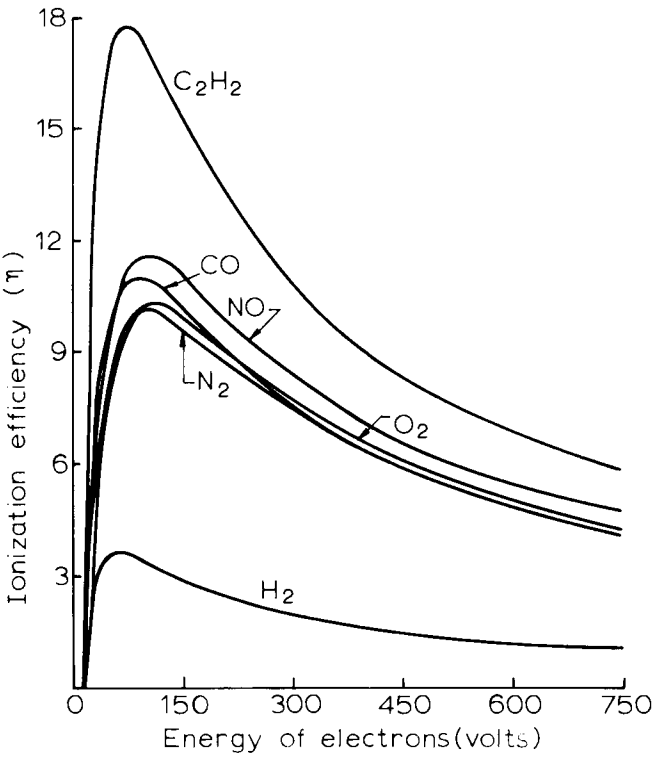


Figure 4.1 The efficiency of ionization of N_2, CO, O_2, NO, H_2 and C_2H_2 according to Tate and Smith[3]

electrode (anode) to draw electron current and a third electrode (collector) to collect the ions formed by electron impact in the gas. Indeed the first ionization gauges were converted triode vacuum tubes, with the grid acting as the anode and the anode used as the collector.

K depends on temperature, as defined above, though it also depends on the gas species, the electron energy and the gauge geometry. The effect of the gas species and electron energy is illustrated in *Figure 4.1* where the number of ions formed per electron per cm path as a function of electron energy is given for several gases commonly encountered in vacuum systems at 10^2 Pa and 0°C. These represent typical ionization efficiency curves applicable to other pressures, although at lower pressures η will be only a fraction of the values shown. Ionization commences at a threshold energy (ionization potential) rises to a maximum and falls again at higher energies. The geometry of the gauge determines the electron path length, the field distribution and ion collection efficiency. A number of workers over the last 50 years have attempted to derive K for particular gauges from fundamental data, but although reasonable agreement between the derived values and measured values has been claimed, in general it is necessary to obtain K from a calibration against a 'standard' (see Section 4.4).

The linear relationship of Equation (4.2) holds from 'zero' pressure up to the pressures where the ion current level affects the electron energy and current by

space charge. However, the upper pressure limit is well above the ultrahigh vacuum range and is not the concern of this book. Although theoretically there is no low-pressure limitation on ion formation, there is the practical problem of measuring these extremely small ion currents. To give an adequate supply of ions according to Equation (4.2), a large electron current and/or a high value of the sensitivity, K, is required. The electron current is limited by power dissipation in the cathode and there are other constraints which will be discussed later. So it is the sensitivity which becomes the important parameter for low-pressure measurements.

In the ultrahigh vacuum region the mean free path of the electrons is very much greater than the linear dimensions of the gauge envelope and the probability of an ionizing collision by an electron traversing between cathode and anode is very small, resulting in a low value of K. As with ion pumps, to increase the ionization efficiency the electron path length has to be considerably extended and this is the main criterion in the basic design of a vacuum ionization gauge. There are several configurations of electrostatic fields, or combinations of electrostatic and magnetic fields, which will cause electrons to oscillate or spiral within a limited volume. Several have been proposed for gauge use but only two or three have evolved into practical ultrahigh vacuum gauges. Although some of the configurations proposed bear resemblance to ion pump designs, the requirements for the collection of ions with low pumping has resulted in rather different solutions. Obtaining sufficient ion current for measurement purposes is not the only problem. At low pressures a more serious limitation is the observed presence of extraneous currents in the gauge which are independent both of pressure and of sensitivity. These result in pressure reading errors since

$$P' = \frac{1}{K}\left(\frac{i_+ + i_s}{i_-}\right) = P\left(1 + \frac{i_s}{i_+}\right) \tag{4.3}$$

where i_s is the total extraneous current, P' the measured pressure and P the true pressure. When $i_+ \gg i_s$, the error is small but as the pressure and thus i_+ decrease, the error becomes increasingly important and eventually limits the lowest pressure measurements, as $P' \rightarrow [(1/K)(i_s/i_-)]$.

The main cause of this extraneous, or residual, current in hot cathode ionization gauges is the so-called X-ray effect. Soft X-rays are produced by electron bombardment of the positive electrode and some of these impinge on the ion collector to produce photo-electrons. The photo-electron current adds to the ion current and cannot easily be separated. In the conventional triode ionization gauge, where the ion collector is a metal cylinder which surrounds the grid anode, the X-ray effect restricts the pressure measurements to about 10^{-5} Pa. Extraneous currents can also be caused by electrical leakage, ion desorption from the anode, again as a result of electron bombardment, photo-emission, caused by radiation from the filament, and field emission. These effects are discussed later when considering the various types of gauge.

Thus, an ultrahigh vacuum ionization gauge must be designed to give as high a sensitivity as possible, subject to having a low value of the residual current i_s. The ionization gauges that have been developed along these lines for ultrahigh vacuum application can be divided into two groups, hot filament (thermionic emission) ionization gauges and cold cathode (field emission or

secondary emission) ionization gauges. In practice the groups are more specifically defined since most of the hot filament gauges are based on the Bayard–Alpert gauge and the cold cathode gauges on the magnetron gauge.

4.2.1 The Bayard–Alpert gauge

Although in the conventional triode gauge the efficiency of collecting the ions was high and a sensitivity of $K = 0.15\ \text{Pa}^{-1}$ for nitrogen could be obtained, the cylindrical collector surrounding the other electrodes also intercepted most of the soft X-rays produced at the grid. This resulted in a high photo-current and thus the severe low-pressure limitation already mentioned. The Bayard–Alpert gauge[1] was primarily designed to reduce the X-ray photo-current by making the collector small so that it did not intercept many of the X-rays. Generally reducing the collector size, however, decreases the ion collection efficiency, for example in the design of Lander[4] the ion current to the collector in the form of a 10 mm diameter disc was about a fifth of that obtained at the cylinder walls. Bayard and Alpert hit on the novel idea of placing as the collector a thin wire, 150 μm diameter, along the axis of the grid structure. Because of the potential 'well' formed by the collector, the majority of the ions formed within the grid are collected and, with suitable dimensions, sensitivities comparable to the conventional gauge were obtained. The basic construction of the gauge is shown in *Figure 4.2*. The thermionic electron emitter in the form of a hairpin or a straight coiled filament is placed outside the grid parallel to the axis. The whole assembly is normally surrounded by a cylindrical screen,

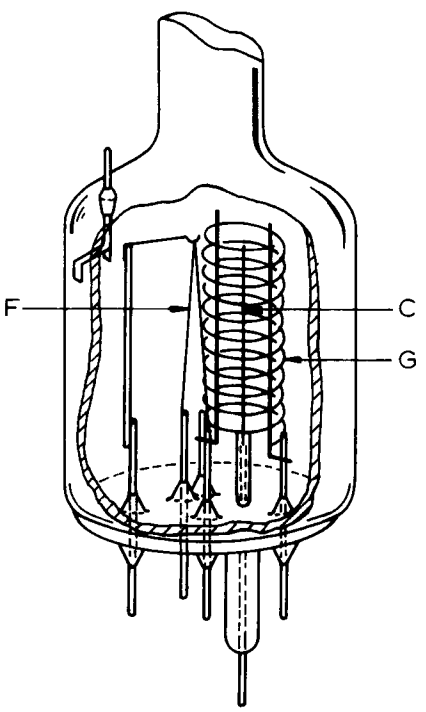


Figure 4.2 Typical commercial Bayard–Alpert ion gauge. F, filament; C, collector; G, grid (anode)

formed either by a tin–oxide coating on the glass envelope or, in the case of a nude gauge*, by a metal screen or by the chamber walls.

The electrons emitted from the filament are accelerated towards the grid by the cathode-grid potential, which is arranged such that an electron entering the grid region has a kinetic energy close to that corresponding to maximum ionization probability of the gas molecules it encounters. Fortunately, as shown in *Figure 4.1*, the required electron energy, 100–150 eV, is much the same for most gases. Because of the logarithmic variation of potential with radius within the grid, most of the potential change occurs near the collector wire and therefore the field is fairly uniform across most of the grid space and the trajectories of the majority of the electrons will suffer little deviation. Having crossed the grid region, the electrons are decelerated in the field between grid and screen and returned towards the grid.

Some typical electron trajectories are shown in *Figure 4.3*. These were computed by Pittaway[5] using programs designed to calculate the potential distribution and charged particle trajectories in electrostatic systems. The

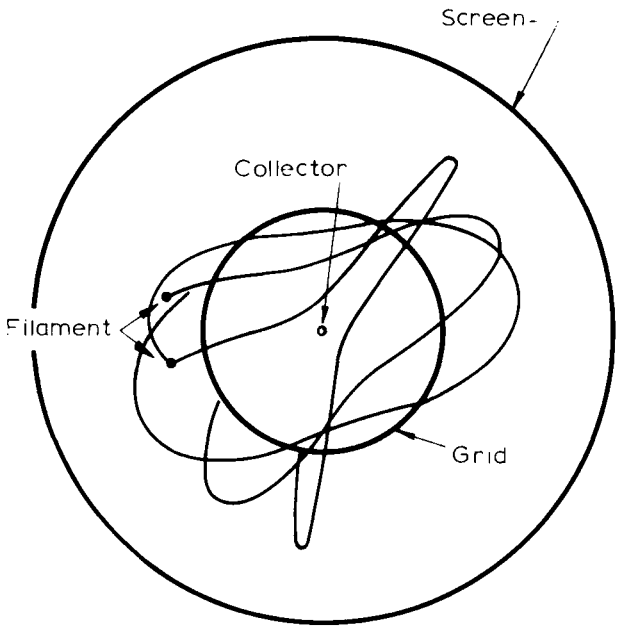


Figure 4.3 Electron trajectories for three different angles of incidence at the grid boundary in a cross section of a Bayard–Alpert gauge according to Pittaway[5]

average number of passes the electrons make through the grid before being collected at the grid wires depends on the grid transparency. According to Pittaway, for the normal grid with a 90% transparency, the average number of passes is five. Since ions are only collected from ionizing collisions within the grid structure, the electron path length within the grid should be as long as

* A nude gauge is the term used for a gauge mounted on a flange without envelope for insertion into the vacuum system

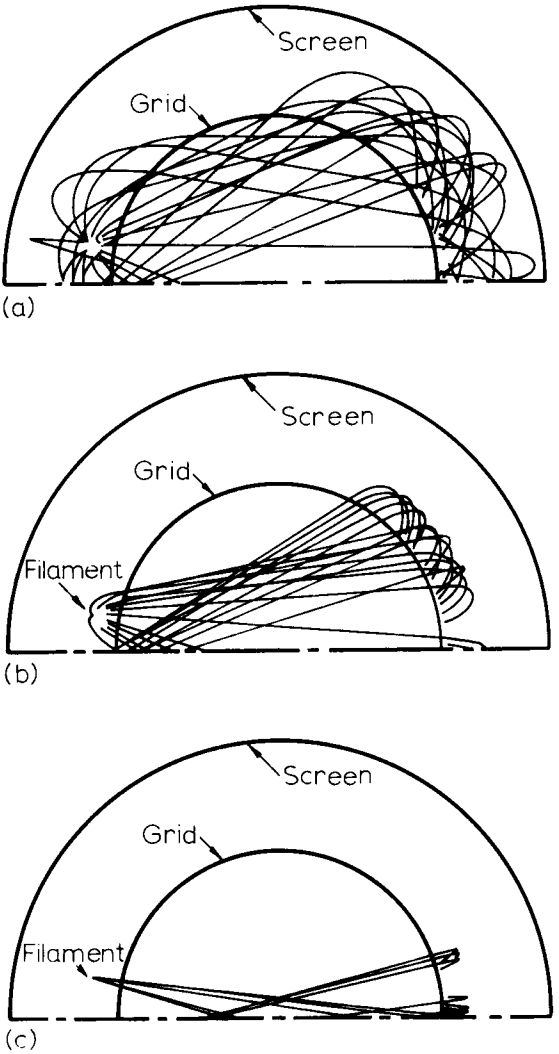


Figure 4.4 Electron trajectories across the grid region of a BAG under different field conditions according to Pittaway[5]. (a) Voltage between grid and screen $V_{gs}=120$ V, Voltage filament to screen $V_{fs}=20$ V; (b) $V_{gs}=250$ V, $V_{fs}=150$ V; (c) $V_{gs}=370$ V, $V_{fs}=270$ V

possible. From *Figure 4.3*, this implies that the angle at which the electrons approach the grid boundary should be as near normal as possible, i.e. in a radial direction. In this respect the screen potential is important. The potential affects the field at the cathode and thus the electron paths, as illustrated in *Figure 4.4* where trajectories are plotted for three different field configurations. Pittaway's calculations confirmed that the variation of K with screen potential found in practice is consistent with the change in electron path length. In early gauges there was no screen and the glass envelope would charge up. As a result

the sensitivity tended to be unstable. Typical values for the electrode potentials required for operating a Bayard–Alpert gauge are

Collector potential	−10 V
Grid potential	180 V
Cathode filament	30 V
Screen	0 (earth)

The optimum values will depend on the gauge geometry and the manufacturer gives recommended operational voltage values.

The early measurements of Bayard and Alpert estimated the residual current to be a factor of 100 below that of a conventional triode gauge, enabling pressures down to 10^{-7} Pa to be measured. Since then the residual currents in the BAG have been the subject of considerable study. It was found that geometry and the condition of the gauge markedly affect the residual current. In a series of experiments with a modulated BAG (see Section 4.2.2), Redhead[6] established two effects: the X-ray effect and electronic desorption of ions from the gas adsorbed on the grid. He showed that the latter contribution to the residual current could be quite large, greater than the X-ray photocurrent if the grid was coated with a considerable layer of gas, especially oxygen. Bombardment of the grid by electrons releases 0^+ ions with energies up to 6 eV. Although at this energy their angular velocity about the collector is sufficient to prevent them reaching the collector under the action of the radial field, those that have no angular velocity or lose it by collision with gas molecules could reach it and thus constitute a significant residual current. Cleaning the grid by heating it with a direct current through the wires, or better still by electron bombardment, drives off the adsorbed gas and reduces the effect significantly. It is generally agreed that provided the desorption effect can be reduced in this way, the residual current in a BAG corresponds to a pressure of 4×10^{-9} Pa. This implies that the lowest pressure which can be measured to an accuracy of 10% is about 4×10^{-8} Pa. An improvement can be obtained in the residual current using a finer collector although this may reduce the sensitivity. Van Oostrom[7], using a 4 μm diameter wire collector, estimated an X-ray limit of 10^{-10} Pa without a similar loss in sensitivity, provided the grid had end caps. The gauge, although used in the Philips Research Laboratories, was not commercially exploited.

An ion gauge strictly measures the gas density within the grid structure. If this is to represent the pressure in the vacuum system then there must be no large temperature variations throughout the system nor must there be any flow of gas to or from the gauge. Although the cathode represents a heat source, the temperature distribution is not normally a problem at low emission currents. The flow of gas, however, can give rise to serious error. It is caused either by gases given off by the gauge or by the gauge acting as a pump. If the gauge is connected to the system by a low conductance, such as a narrow tubulation, this gas flow can give rise to quite large pressure differentials across it. To reduce the gas evolution, the gauge should be bakable to at least 250°C and the components, which should consist of the minimum amount of metal, should be capable of being heated to red heat. As already stated, it is particularly important to clean thoroughly and outgas the grid structure to prevent ion desorption currents.

Pumping of gas in an ionization gauge is due to two effects, ion pumping and

chemical pumping. Ion pumping is the trapping of the energetic positive ions which become imbedded in the surface on which they impinge, i.e. either at the collector or screen or walls of the gauge. Chemical pumping is caused when active gases react with gauge components, particularly with the heated filament. As we saw in Chapter 3, pumping by the gauge can be quite high, of the order of 10^{-4} m^3s^{-1} and Alpert[8], in his experiments on ultrahigh vacuum in the 1950s, advocated their use to pump down small systems. Since the pumping is inherent in the mechanism of the gauge it cannot be entirely eliminated. It can however be minimized by using the lowest electron current consistent with a measurable ion current. For most gauges this is around 100 μA but it has to be raised at the lowest pressures.

The minimum ion current is of the order of 10^{-13} A and in order to measure such currents it is essential that the electron current is stabilized. The cathode operates in a temperature limited mode, so that electron emission is controlled by regulation of the cathode temperature. This is achieved by a feedback system whereby the electron current is sensed and used to control the filament supply and thus its temperature. In parallel with the development of gauges there has been a continual improvement in the control and measuring circuits. Initially thermionic valve circuits were employed but transistor and integrated circuits have now been introduced.

Figure 4.5 shows a basic circuit using an integrated differential amplifier[9]. The electron accelerator voltage is applied between grid and cathode via a current sensing resistor R_1. A voltage proportional to the cathode current is developed across the resistor R_1 and compared with a reference voltage by the amplifier A. Any out-of-balance voltage is amplified and added to or subtracted from the filament supply. The reference voltage is set at the required ion collecting bias potential. In this way the amplifier performs two functions in that it sets the correct cathode current and also the ion collector bias potential automatically.

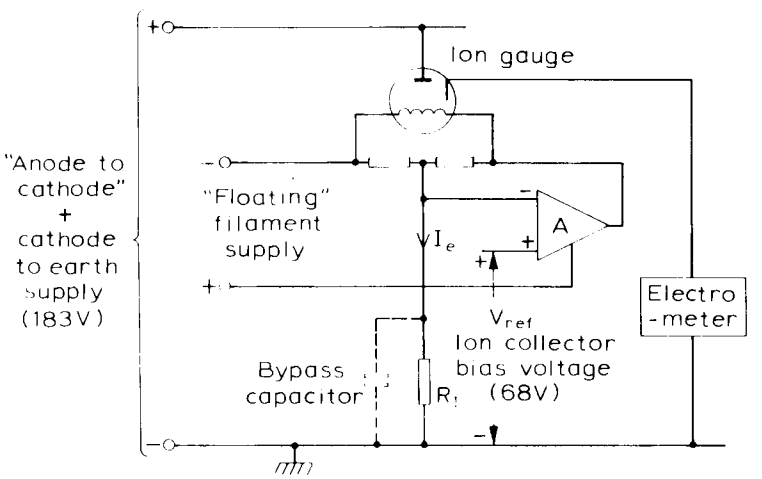


Figure 4.5 Block diagram of emission control circuit using an integrated operational amplifier, based on circuit of Close and Yarwood[9]

Herbert[10] has described an alternative method whereby the collected electron current rather than the emission current is stabilized. This is useful in more complex ionization gauges and mass spectrometers where more than one grid or anode is used but where the total emission current may not represent the true ionizing current. Such circuits are capable of holding the ionizing electron current stable to within 1% over a current range of at least two orders of magnitude. Normally the control unit will also include protection circuits and facilities for outgassing the gauge. Protecting the filament, by extinguishing it if the pressure rises above a fixed value, is necessary because the tungsten filament commonly employed will rapidly oxidize and evaporate at high oxygen ambients. Even at vacuum pressures below 10^{-4} Pa thinning of the filament by oxidation and/or evaporation represents the main limitation on gauge life, although 5000–10 000 h can be expected with careful use. Some gauges contain a second cathode filament to be used when the first fails and nude gauges commonly have facilities for changing filaments. If oxygen pressures are to be measured then it is possible to use a filament with a lower work function that can be run at a lower temperature. A rhenium filament coated with lanthanum hexaboride[11] is suitable for such applications. Coated filaments, however, must be used with caution since evaporation of the low work function material on to the grid can significantly increase the extraneous current.

As a result of the considerable studies that have been made on the Bayard–Alpert gauge, it has become established as the universal gauge not only for ultrahigh vacuum down to 10^{-8} Pa but also for high vacuum covering the pressure range of 10^{-1} Pa to 10^{-5} Pa. Designs are commercially available from most vacuum equipment companies, together with suitable control supplies, and provided the gauges are properly outgassed, etc. reliable measurement down to pressures approaching 10^{-8} Pa can be expected from them.

4.2.2 The modulated Bayard–Alpert gauge

The modulated BAG was introduced by Redhead in 1960[12] to extend the range of the BAG to lower pressures. It consists of a standard BAG with the simple addition of a fine wire electrode mounted parallel with the grid axis inside the grid structure, *Figure 4.6*. With the modulator at grid potential the current to the collector is unaffected and the collector current is given by

$$I_1 = i_+ + i_s \tag{4.4}$$

If now the modulator is biased to the collector potential it will draw a fraction of the collector current αi_+, whereas the X-ray photo-current should be unaffected. The new collector current will be

$$I_2 = (1-\alpha)i_+ + i_s$$

and thus

$$I_1 - I_2 = \alpha i_+ \tag{4.5}$$

where α is termed the modulation factor and can be determined at higher pressures where the residual current, i_s, is negligible compared with i_+. The value of α is approximately 0.6 for a typical modulated BAG.

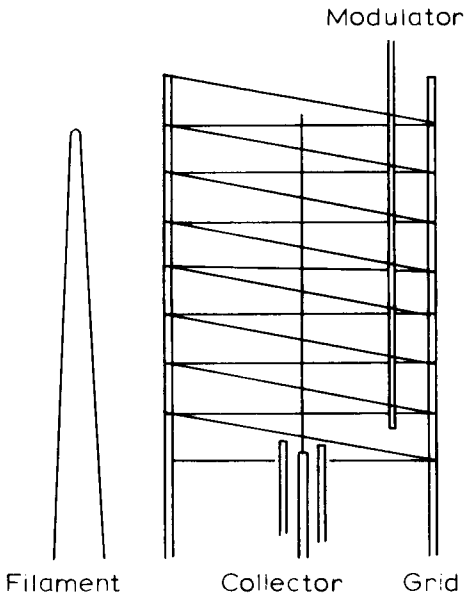


Figure 4.6 Schematic diagram of a modulated Bayard–Alpert gauge

The validity of the modulator gauge depends on the assumption that the residual current is unaffected by the potential on the modulator. The steep potential well around the collector will hardly be affected by the modulator potential and therefore the photo-electron current, which would only be expected to alter if the field changed at the collector, should not be modulated. Redhead[6] also claimed that the desorbed ion current from any gas layers on the grid should also remain unmodulated. The reason for this is that the desorbed ions are released with a high mean kinetic energy, typically 5–6 eV for 0^+ ions from oxygen contaminated molybdenum, relative to that of the gas-phase ions. As a result the influence of the modulator potential on their trajectories would be small. In effect, the desorbed ions act in a similar way to X-rays and only those directed towards the collector will reach it. Nevertheless studies of the modulated BAG at very low pressures, for example by Hobson[13], produced evidence that the residual current itself had a modulation factor which restricted the pressure measurements to above 4×10^{-9} Pa.

Lange and Singleton[14] also pointed out that outgassing of the modulator wire could be a problem. At collector potential, gas ions are collected but at grid potential, the wire is bombarded by electrons which would release the ion-pumped gas. They advocated modulating over a much smaller voltage range (grid voltage, V_g, to $V_g - 20$ V) to reduce this effect. This appeared satisfactory with the standard open ended grid structure of *Figure 4.6*. However, it has been found that closing the grid at each end with a mesh through which the collector and modulator pass gives better reproducibility and stability and is preferred for a modulator gauge. For such a structure the depth of modulation advocated by Lange and Singleton is rather low.

Van Oostrom[15] found that he could use the end caps of his fine wire gauge

as the modulator. Only a small change in the potential distribution was needed to alter drastically the collection efficiency of the gas ions without affecting the faster desorbed ions. A modulation factor of 0.9 was achieved. The system would not be satisfactory with the normal 150 μm diameter collector since a greater change in the field distribution would be required, i.e. a larger swing of the voltage on the caps and this would cause the electron current to be modulated.

The reason for the modulation of the residual current components is not clearly understood. There are several related factors, for example the electron current to the modulator, when at grid potential, is not insignificant but may not be included in the emission control circuit, so that the electron current is modulated. Also, by using a gauge which was heavily contaminated with adsorbed gas, Pittaway[16] has demonstrated that, although desorbed ions from the grid may not be modulated, desorbed ions from the modulator itself could be focused on to the collector to give an enhanced collection factor for desorbed ions at certain field configurations. It is interesting to note that this deduction was made following the observation that the ion current increased instead of decreasing when the contaminated BAG was modulated, a phenomenon also observed by other workers. In view of these uncertainties the modulated BAG must be used with caution. However, it does extend the range of the BAG by an order with probably less than a 10% error for a 'clean' gauge and it has become established as a practical gauge for pressures below 10^{-8} Pa.

Normally the modulation is carried out manually but proposals have been made to apply a.c. signals to the modulator and automatically record the pressure[17,18]. The main problem with this technique is the capacitive current which has to be compensated or allowed for.

The modulation technique can also be applied to other gauges such as the extractor gauges, described in the next section, to extend their range.

4.2.3 Extractor gauges

The basic principle of the extractor gauge is to locate the collector outside the grid structure of the gauge in which the ions are formed, so that X-rays cannot reach it. Sufficient ion collection is maintained by 'extracting' ions from the grid region into the collection region. In this way the residual current is reduced allowing lower pressures to be measured. In most designs the ions are extracted through an aperture in an end-cap or screen across the end of the grid to an externally mounted collector. In some designs the end-cap, through which the ions are extracted, is separately connected and biased below the grid potential to accelerate the ions from the grid cage. In others the end-cap is connected to the grid and extraction depends on a field penetrating through the aperture from an external electrode. Whichever method is used, the X-rays reaching the collector are restricted by the end-cap to those having direct line of sight to the collector through the aperture.

The first design of gauge having an external collector was described by Schuemann in 1963[19] and is illustrated in *Figure 4.7*. However, although the grid end-cap extracted the ions and restricted the X-rays from the grid structure, the collector plate was large and the X-ray current was of the same order as for a BAG. The gauge was not primarily designed as an extractor

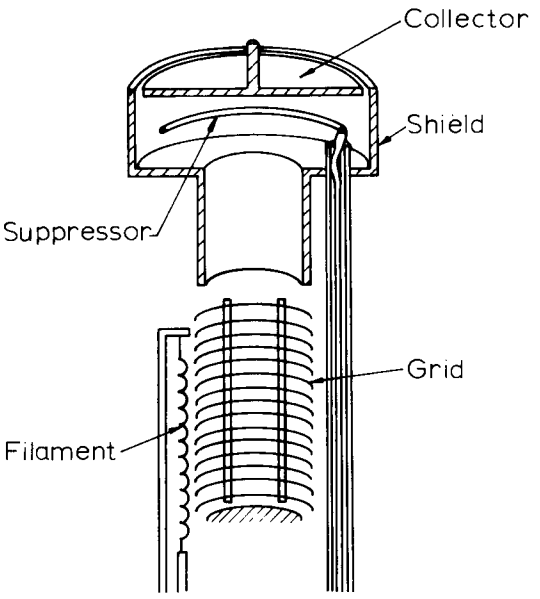


Figure 4.7 Schematic diagram of the screened collector gauge of Schuemann[19]

gauge and the improved performance was achieved by the addition of a suppressor ring which was negatively biased to prevent photo-electrons leaving the collector. The idea of a suppressor electrode was first suggested by Metson in 1951[20] who inserted a second grid as suppressor in the standard triode gauge which extended its range down to 10^{-7} Pa. Metson's suppressor gauge was limited because the suppressor itself became a source of photo-electrons, and since it was biased negative to the collector, the electrons emitted could reach the collector. In the Schuemann gauge the suppressor ring was screened from the incident X-rays and pressures down to 10^{-9} Pa could be measured.

The name extractor gauge was coined by Redhead[21] in 1966 with the gauge shown in *Figure 4.8*. The collector consisted of a short fine wire mounted opposite the aperture in the end shield about 12 mm from the end of the grid cage and surrounded by a hemispherical ion-reflector held at grid potential. The shield was negatively biased to extract the ions. The gauge also incorporated a modulator to assess the X-ray limit. The sensitivity of the gauge was approximately $K = 0.09\ \text{Pa}^{-1}$ for nitrogen but depended on electrode potentials and the electron current. The residual current due to X-rays was calculated by considerations of geometry to be better than for the BAG by a factor of 160. It was not measurable as a change in modulation factor down to the lowest pressure attained of around 10^{-10} Pa. Redhead also found that desorption ion current from a contaminated grid was very much reduced and assumed that this was due to the low efficiency in collecting the energetic ions emitted from the grid. This has been confirmed by Pittaway[22] with plots of ion trajectories for gas phase and desorbed ions in the Redhead gauge geometry.

At about the same time Groszkowski[23,24] was examining the effect of

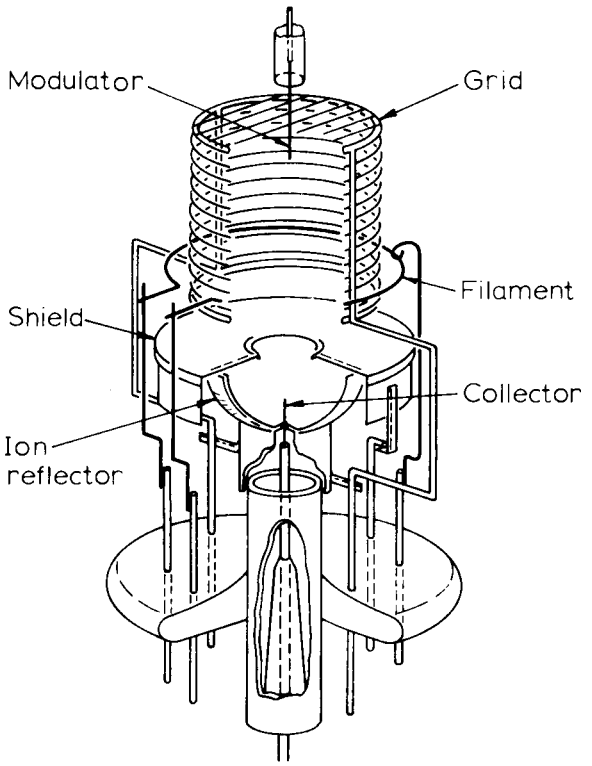


Figure 4.8 Cutaway diagram of Redhead's extractor gauge[21]

withdrawing the collector from a BAG along the axis. He showed that a reduction in the X-ray limit of about two orders of magnitude could be achieved if the position of the collector was such that only X-rays from the far end-cap could reach it. The design of gauge which he developed is shown in *Figure 4.9(a)*. He used a thin, short collector placed inside a glass sleeve, located outside the grid box opposite an orifice in an anode end-cap which was electrically connected to the grid. Extraction in this case depends on penetration of the field into the grid cage and equipotentials and ion trajectories obtained by Pittaway[22] using a solid end cap and Groszkowski's geometry are illustrated in *Figure 4.9(b)*. Results on this gauge seemed to be rather variable and, in particular, Pittaway found the sensitivity to be critically dependent on the electrode potentials and electron current. He explained this in terms of the potential distribution inside the grid. For a normal helical grid the potential inside the grid gradually falls with increasing radius to a value some volts below the grid potential, due to field penetration through the grid wires. In particular the penetration is dependent on the potential difference between the grid and the glass walls of the envelope around the gauge. Since these glass walls can charge up, a variation of potential fall off could be expected. If the potential within the grid was as much as 20 V below the grid potential, which was the case when the cathode filament, and hence the charged glass walls were 200 V below grid potential, then the ions which

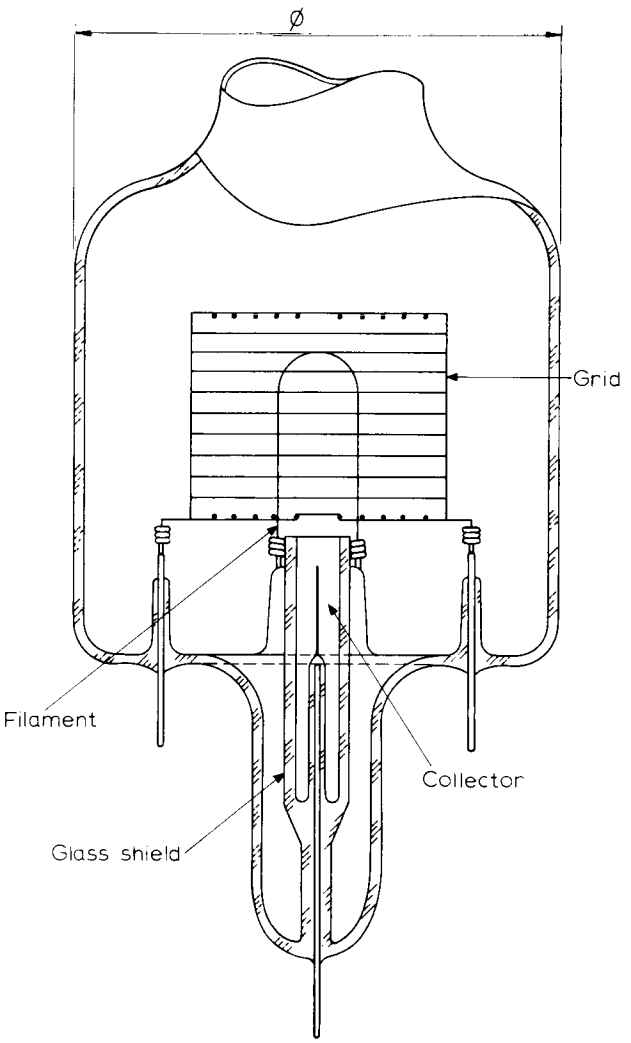


Figure 4.9 (a) Schematic diagram of the extractor gauge of Groszkowski[23];

should have been trapped in the grid cage could escape between the grid wires rather than through the aperture. A vast improvement was obtained by using a grid with very small mesh apertures to inhibit field penetration from the filament and walls whilst still maintaining a high electron transparency. Pittaway[25] achieved this with a finely woven tungsten mesh. The effect on sensitivity is shown in *Figure 4.10.*

From his studies Pittaway designed an extractor gauge which combined the attributes of other designs with a fairly simple structure which could be operated from a standard BAG control box. The design is shown in *Figure 4.11* together with a diagram showing the field distribution and some ion trajectories for normal operational conditions. The grid end-cap screen was

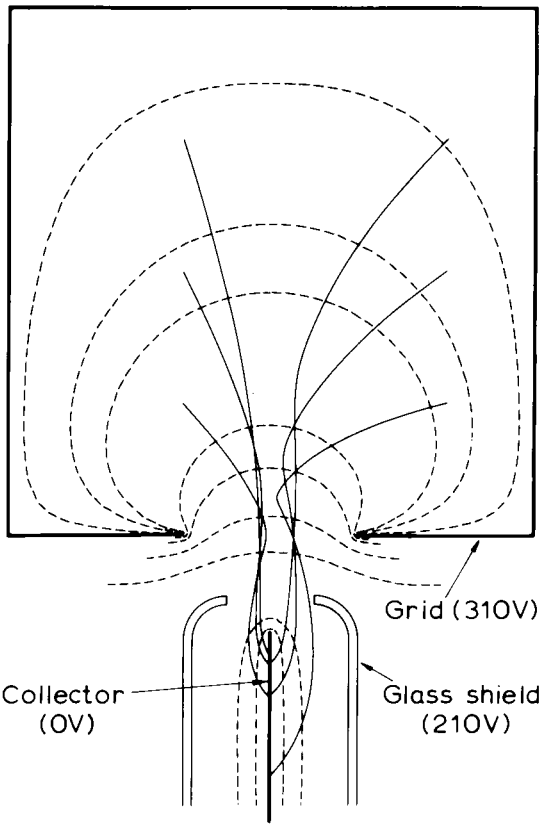


Figure 4.9 (b) Field distribution and ion trajectories in the Groszkowski gauge according to Pittaway[22]

electrically connected to the grid and a separate extractor electrode was used to draw out the ions and at the same time to provide a field barrier to energetic electrons. The collector was surrounded by a reflector which focused the ions on to the collector but was shaped with an inner cone to trap reflected X-rays. The glass bead was intended to shield the thicker collector support rod from incident X-rays. The potentials on the electrodes and their positioning formed a lens system which prevented ions from reaching the extractor and focused the majority on to the collector. The sensitivity was around $K = 0.09\ \text{Pa}^{-1}$ for nitrogen and was reasonably constant with electron current and energy. The aperture in the end-cap was only 4 mm diameter so that the source of X-rays which were incident on the collector was limited to a small area of the grid cage. The estimated X-ray limit corresponded to a pressure of 10^{-10} Pa. By incorporating a modulator electrode in the form of a short wire at the other end of the grid cage a high modulation factor of 0.95 was achieved. Using the modulator it was claimed that pressures down to 10^{-12} Pa should be measurable.

By modulating the ion current in the region between the ionizing area and the collector, Watanabe *et al.*[26] claim that the photo-current due to X-ray

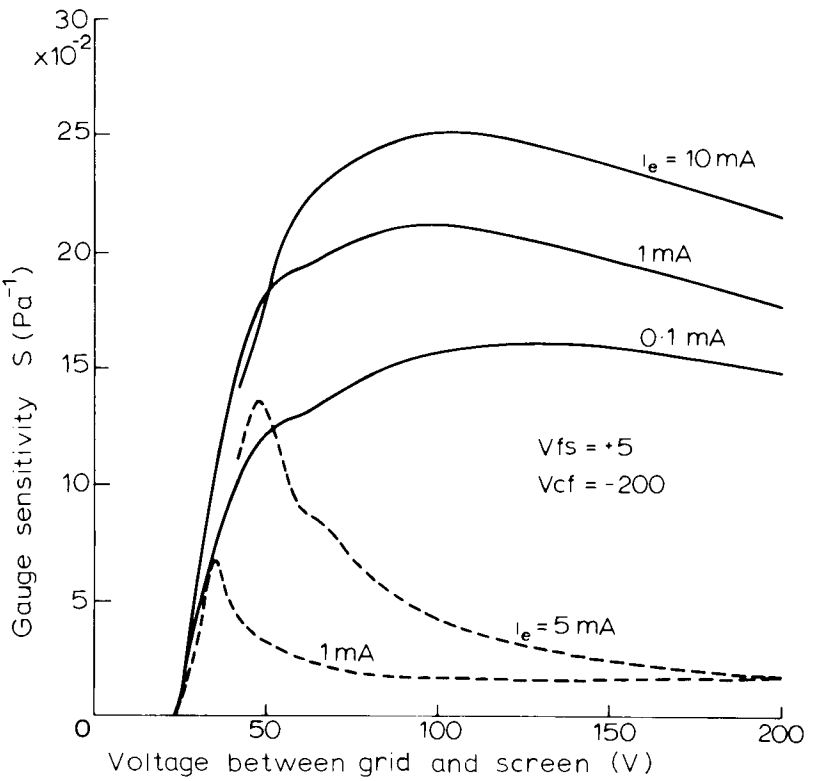
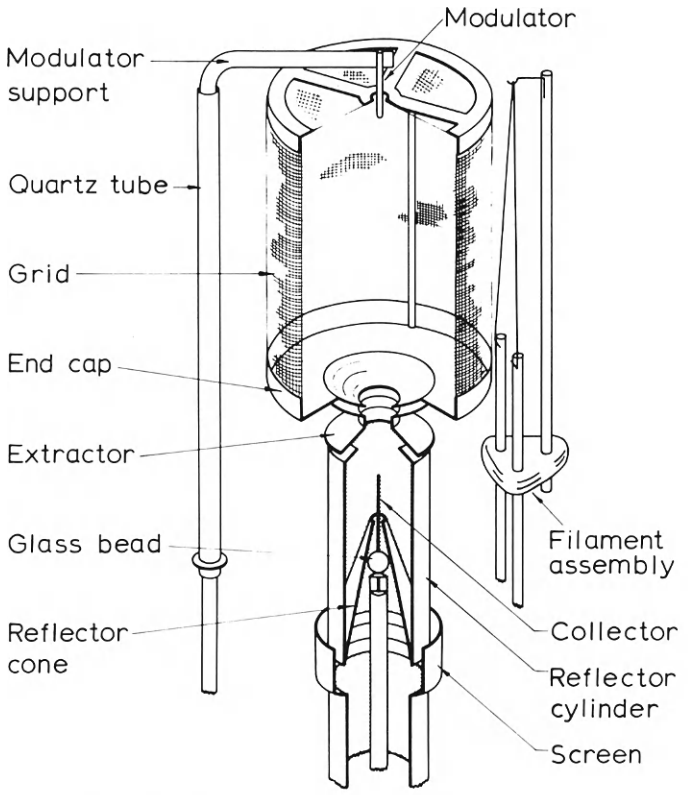


Figure 4.10 Variation of sensitivity with grid potential V_{gs} and electron current i_e in an extractor gauge with a fine mesh structure. Corresponding curves for the Groszkowski gauge are shown as dashed curves. Pittaway[22]

radiation will not be modulated and would not represent the limit in an extractor gauge. The construction of their gauge was designed to work in the noisy environment of high energy accelerators and was rather different from the normal extractor gauge. The head had a hemispherical mesh anode with a funnel-shaped electrode across its diameter through which the ions were accelerated. A cylindrical electrode between the ionizing area and the collector was used to modulate the ion current. By sine wave modulating at 12.4 Hz the output could be a.c. coupled to the amplifier and phase sensitive detector. Pressures down to 10^{-8} Pa were measured.

A rather different type of extractor gauge was described by Helmer and Hayward in 1966[27]. In their gauge the ions were extracted with a potential on the exit slit and bent through 90° with an electrostatic field. The basic arrangement is shown schematically in *Figure 4.12*. In this way X-rays can only reach the collector plate by reflection. As an added precaution a suppressor grid was included. Although the suppressor did not appear necessary as far as X-ray photo-currents were concerned, it could be used to examine the ion energies and to show that part of the ion current was due to



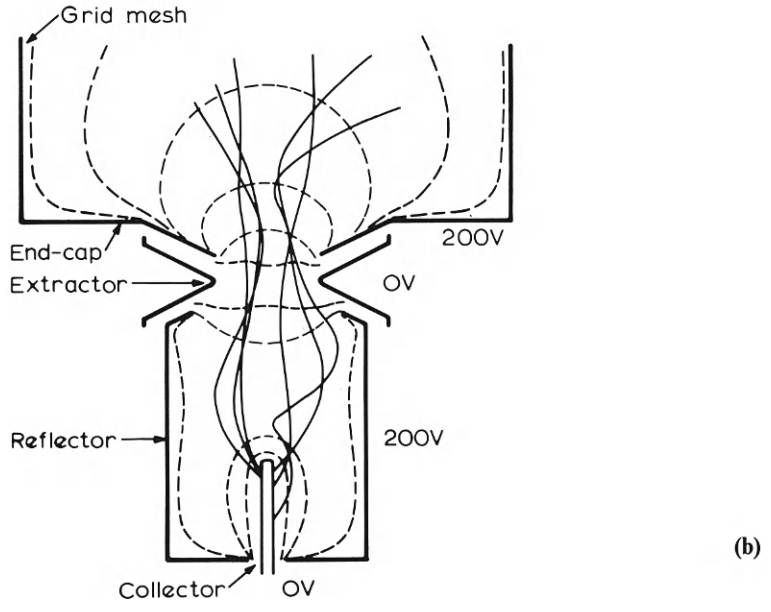


Figure 4.11 (a) Cutaway diagram of the extractor gauge of Pittaway[25]; (b) Equipotentials and ion trajectories in the gauge without the reflector cone

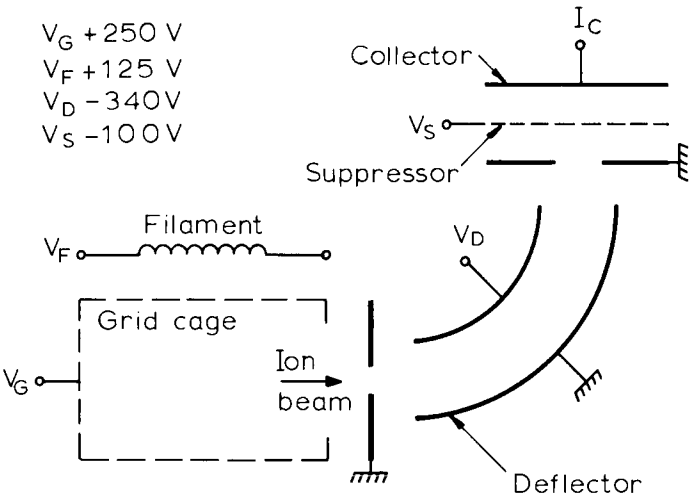


Figure 4.12 Schematic diagram of the bent beam gauge of Helmer and Hayward[27]

secondary ions produced by ions bombarding the electrodes of the deflection system. However, since these secondary ions were also pressure dependent, they did not represent an extraneous current to affect the measurements. The sensitivity of the gauge was around 0.12 Pa^{-1} for nitrogen but it was pointed out that the design allowed for incorporation of an electron multiplier which would increase the sensitivity. Residual currents equivalent to a pressure of 2×10^{-12} Pa were indicated, suggesting that the X-ray limit was probably around this figure. The gauge is rather more complex in structure with more metal to outgas than in other designs and requires additional voltage supplies.

Summing up, extractor gauges offer an improved performance over the BAG with a lower pressure measuring capability. However, they are more complex in construction and some are more difficult to use. This, coupled with their more limited application has made them expensive gauges. They have therefore had limited commercial exploitation and few vacuum equipment manufacturers include them in their range.

4.2.4 The magnetron and similar gauges

The magnetron type gauge tackles the problem of residual current limitation by the alternative approach of increasing the sensitivity by an order or more. This is achieved by introducing a magnetic field which causes the electrons to take long spiral paths between electrodes, thus increasing the probability of ionizing collisions.

The principle of using a magnetic field crossed with an electrostatic field to obtain increased ionization at low pressures dates from 1937, with the Penning

manometer named after its inventor[28]. The original design consisted of two parallel plate cathodes with a ring anode between them with its axis perpendicular to the plates. A magnetic field of around 400 gauss was applied perpendicular to the cathode plates. Electrons were emitted from the cold-cathodes by ion bombardment, as in a glow discharge and were accelerated to the anode across a potential of around 2 kV. Because of the magnetic field, the electrons travelled in long helical paths of several hundred times the direct distance before reaching the anode. The total current passing, electron plus ion current, was taken as a measure of the pressure and the gauge covered a range of 1 to 10^{-3} Pa. Below this pressure the discharge would extinguish or fail to ignite. An improved version using a cylindrical anode was described by Penning and Nienhuis[29] in 1949 which extended the range down to 10^{-5} Pa (see *Figure 4.13*).

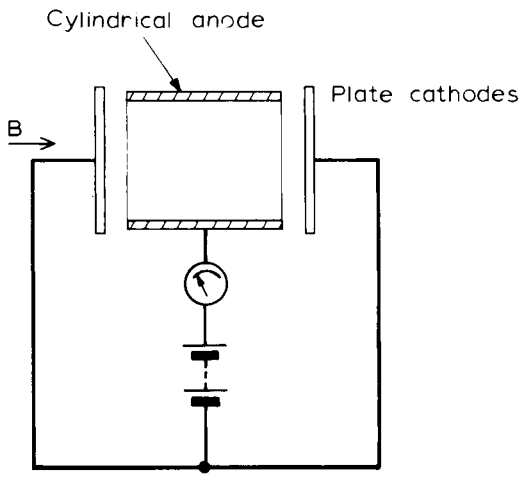


Figure 4.13 Schematic diagram of the basic Penning cell gauge

This type of gauge is not limited by the X-ray effect since the electron current producing the X-rays is pressure dependent, i.e. the residual X-ray photo-current is a function of pressure. However, the high field at the cathode particularly at the edges can cause field emission which will be independent of pressure and this will introduce a low pressure limitation analogous to the X-ray effect in the BAG.

To overcome the problems of extinction of the discharge at low pressures and the field emission, Hobson and Redhead[30] introduced the 'inverted magnetron gauge' in 1958. The basic structure is shown in *Figure 4.14*. It is a three-electrode device consisting of an anode, cathode as ion collector and an auxiliary cathode. The ion collector is a closed ended cylinder with centre holes in the end plates through which the rod-anode and shielding tubes pass. The latter are connected to the auxiliary cathode which is a box-like structure

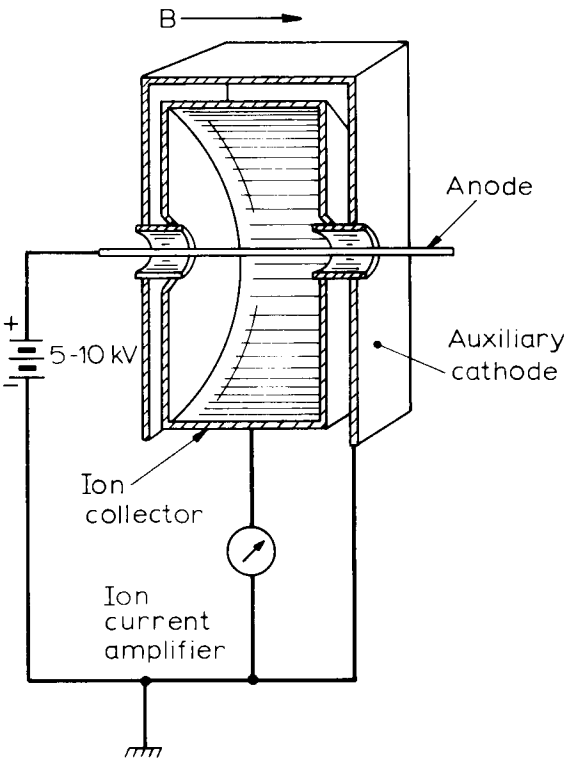


Figure 4.14 Cutaway diagram of the inverted magnetron gauge of Hobson and Redhead[30]

enclosing the ion collector and acting as an electro-static shield to prevent field emission from the collector. The magnetic field is applied along the axis. Using 6 kV on the anode and a magnetic field of 2060 gauss, the discharge could be struck at pressures below 10^{-8} Pa, with evidence of pressures being measured down to 10^{-10} Pa. The gauge was subject to oscillations, and the calibration was non linear, $i_c = cP^n$ where n varied between 1.1 and 1.4 for different gauges.

About a year later Redhead[31] described a magnetron gauge to overcome some of the problems with the inverted magnetron. This was similar to a Penning cell with an axial rod connecting the two cathode plates, i.e. the geometry of a normal magnetron. The actual gauge configuration is shown in *Figure 4.15*. The auxiliary cathodes were inserted between the cathode plates and the anode cylinder to reduce field emission. The perforations in the anode were to improve gas flow in the gauge. The magnetron gauge has the advantages over the inverted magnetron of a linear relationship between ion current and pressure over most of the pressure range and also a higher sensitivity. The sensitivity was in fact approximately 45 times greater than that of a BAG.

The electrons released from the cathode orbit the cathode rod with a cyclic motion. The size of their orbit depends on the electron energy and, if the electrons lose energy by collision with gas molecules then they will take up orbits closer to the anode. Thus electrons can reach the anode only by

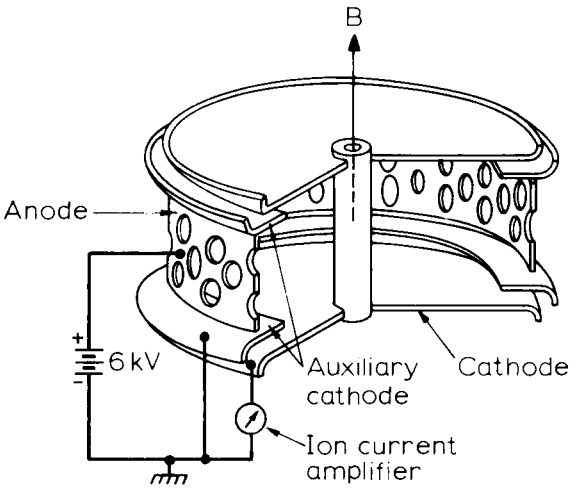


Figure 4.15 Cutaway diagram of the magnetron gauge of Redhead[31]

successive collisions with gas molecules. Unfortunately, the higher ion current, implicit in the high sensitivity, results in a faster pumping speed. This is the main disadvantage of the crossed-field, cold-cathode gauge where pump speeds of the order of 10^{-3} m^3s^{-1} (an order greater than for the BAG) are quite common. Further, because of the cold-cathode discharge mechanism, the current cannot be controlled in the same way as in the BAG and stability can be a problem.

Kageyama *et al.*[32] have suggested a method of operating the gauge to reduce the pumping effect. They apply a step voltage to the gauge and, when the current reaches a saturation value, switch off the supply. The gauge is thus pulsed at a low duty cycle reducing the ion current and pumping action. The pressure is determined by measuring the saturation current and/or the time taken to reach saturation. Pressures were measured over a range of 10^{-6} to 10^{-1} Pa with 10% accuracy over the range but with a 0.6% accuracy around 10^{-4} Pa.

The crossed-field gauge can also be used with a heated filament to supply an electron current. Young and Hession[33] and also Nichiporovich and Khanina[34] have incorporated filaments purely to provide adequate electrons to start the discharge. In this way ignition is assured at low pressures, thus extending the pressure range. These gauges are often referred to as triggered discharge gauges. The full potential of a thermionic emitter in a crossed-field gauge was exploited by Lafferty[35] in his 'hot cathode magnetron'. The initial design is shown in *Figure 4.16*, but modifications and improvements have been described since then[36]. The hair-pin thermionic filament is mounted on the axis of the cylindrical anode. Escape of electrons is prevented by two negatively biased end shields one of which is used to collect the ions. The axial magnetic field is provided by a cylindrical magnet mounted around the bulb. Although collection of ions at only one of the plates reduces the sensitivity, the other plate does not have to be so carefully isolated and this facilitates easier mounting. The gauge has a high sensitivity of the order of 5×10^4 Pa^{-1} which can be improved further by use of a multiplier.

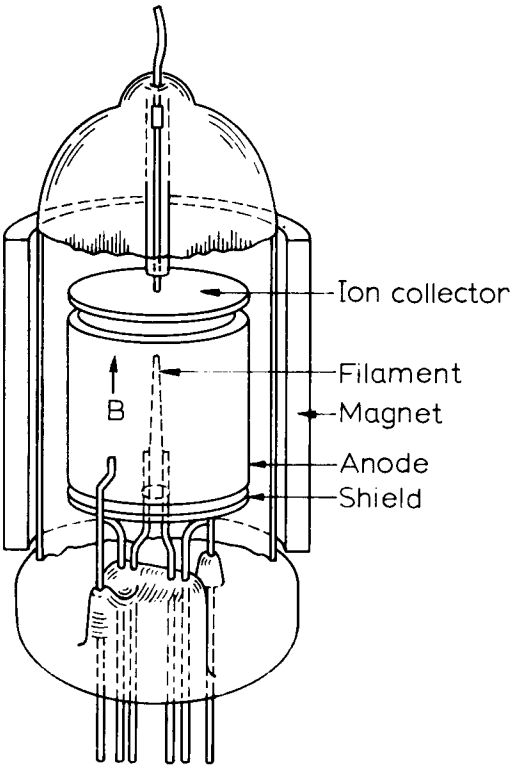


Figure 4.16 The hot cathode magnetron gauge of Lafferty[35]

The current can, to some extent, be controlled, although in practice an electron current of less than 10^{-7} A is required to prevent spurious oscillations. However, a low current has advantages. First, the cathode can be run at a lower temperature to minimize cathode–gas interaction; secondly the low current reduces the ion-pumping effect to about 2×10^{-5} m^3 s^{-1} and thirdly the X-ray effect is small. The latter has been calculated to be equivalent to a pressure of around 10^{-12} Pa. The anode voltage requirement is $\sim$300 V with the cut-off magnetic field being a few hundred gauss. Thus field emission as a limitation can be discounted. Below 10^{-6} Pa the ion collector current is a linear function of pressure. The gauge therefore has most of the attributes one could wish for on ultrahigh vacuum gauge. Its main limitation occurs at pressures of 10^{-7} Pa and above. At such pressures the electron current is pressure dependent and therefore cannot be continuously regulated as with the BAG.

A gauge which should not strictly come in a section on a crossed-field gauge but which has similar properties, is the orbitron gauge[37]. In this gauge the electron paths are extended by constraining them to orbit in an electrostatic field set up between two concentric cylinders. The basic structure is shown in *Figure 4.17*. Electrons from the filament are ejected into the electrostatic field between the outer cylinder at 0 V and the axial anode wire at

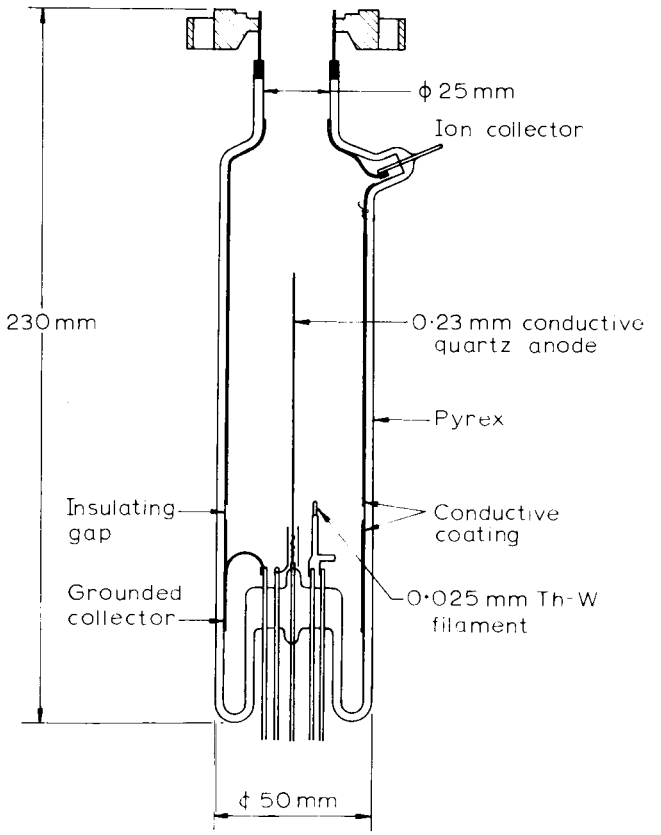


Figure 4.17 Schematic diagram of the orbitron gauge of Meyer and Herb[37]

+500 V which also collects the electron current. The outer cylinder collects the ions formed in the annular space which is taken as a measure of the gas density. A sensitivity of 7×10^2 Pa^{-1} is attained which allows the gauge to be run at a low electron current, 10^{-6} A. The X-ray limit is found to be equivalent to a pressure reading of 10^{-11} Pa. The energy of the electrons entering the gauge depends on the filament bias and position and, since this affects the path length, sensitivity is rather critically dependent on these parameters[38]. The orbitron has not been further developed as a gauge but the application of the orbitron principal as a pump has been given more attention (see Section 3.7).

In general, magnetron type gauges offer the potential of measuring pressures below the limit of a BAG, namely below 10^{-8} Pa. The cold cathode designs are the simplest but have the disadvantage of a relatively high pumping speed and difficulty in stabilizing the current. The hot cathode magnetron has a much lower pumping speed and because the emission current can be low, interaction with active gases, desorption and X-ray effects are less than for the BAG. It appears to have all the advantages required for a low-pressure gauge, but unfortunately there are problems in measuring pressures above 10^{-6} Pa, thus limiting its use. Also the need for a magnet means they

cannot be used as 'nude' gauges. The orbitron has the advantage of no magnetic field, but has not been further developed as a gauge. Since there are a limited number of applications requiring pressures below the BAG range, the commercial development of such gauges has been limited.

4.3 Momentum transfer gauges

Molecules hitting a surface will rebound with a momentum mv where v is the rebound velocity which will be greater than the average molecular momentum of the rest of the gas if the surface is moving or is above ambient temperature. The total momentum of molecules leaving a unit area of surface per second will then be mvv where v is the rate of incidence of molecules on unit area, which, according to kinetic theory (see Chapter 1), can be related to pressure and temperature by the equation

$$v = P(2\pi mkT)^{-1/2} \tag{4.6}$$

If we consider two surfaces, designated 1 and 2, placed distance, d, apart such that d is much less than the mean free path of the molecules, then the rate of momentum transfer, B, from one to the other by the molecules is

$$\begin{aligned} B &= mv(v_1 - v_2) \\ &= P\left(\frac{m}{2\pi kT}\right)^{1/2}(v_1 - v_2) \end{aligned} \tag{4.7}$$

Thus, the transfer of momentum is linearly dependent on pressure with, theoretically, no lower limit.

For the case of one surface moving with a velocity, u, relative to the other, Equation (4.7) reduces to

$$B = P\left(\frac{m}{2\pi kT}\right)^{1/2} u \tag{4.8}$$

On the other hand, if one surface is at temperature T_1, greater than the ambient temperature T_0, then applying kinetic theory for the rebound velocities

$$\begin{aligned} B &= P\left(\frac{m}{2\pi kT_0}\right)^{1/2}\left\{\left(\frac{3kT_1}{m}\right)^{1/2} - \left(\frac{3kT_0}{m}\right)^{1/2}\right\} \\ B &= P\left(\frac{3}{2\pi}\right)^{1/2}\left\{\left(\frac{T_1}{T_0}\right)^{1/2} - 1\right\} \end{aligned} \tag{4.9}$$

It will be noted in this case that the transfer of momentum is independent of the mass of the molecules.

The rate of transfer of momentum will result in a force on the surfaces and it is the measurement of this force which is used as a pressure sensor. At low pressures the force involved is very small and to be measurable, the work done must be integrated over a sufficient time interval and the other forces acting on the surface must be small in comparison. Over the years since the principles were first exploited in 1900, two types of gauge have been developed: those depending on a moving surface, the so-called viscosity gauges, and those

depending on temperature differences, the so-called radiometer gauges. Numerous configurations for such gauges have been proposed, although as one designer puts it, 'throughout their long history there have been more designed and built than used'.

Taking the viscosity gauge first, two different methods of applying the principle have been examined. In the first, the damping effect of the momentum transferred to a vibrating surface is measured, usually by setting the surface oscillating and measuring the decay time. In the second method a surface is continuously rotated and the torque is measured either to a second stationary surface placed adjacent to the rotating surface or by the drag on the rotating surface. Early designs are described by Dushman[39]. In general the lowest pressures recorded were in the region of 10^{-4} Pa and at such pressures delicate suspensions were required to ensure that friction or stiffness of the suspension did not mask the force due to momentum transfer. The breakthrough to a more rugged and reliable gauge capable of measuring lower pressures came with the idea of using a rotating ball or disc freely suspended in a magnetic field. The effect of the vacuum pressure on a spinning ball bearing levitated in an electromagnetic field was first mentioned by Beams *et al.* in 1946[40], but it was not until 1962 that the idea was investigated as a vacuum gauge[41]. The ball was suspended in a vertical axial magnetic field provided by an electromagnet which was regulated by the vertical position of the ball via a sensor and servo-mechanism. The ball was rotated up to a speed of 100 000 rev/s and then allowed to coast. The decay of speed with time was measured. If D is the tangential damping force due to momentum transfer between the gas molecules and the ball then

$$-I\frac{\mathrm{d}\omega}{\mathrm{d}t}=D\omega \tag{4.10}$$

where I is the moment of inertia of the ball and ω the angular velocity in radians per second. If ω is expressed in terms of revolutions per second, N, then $\omega=2\pi N$ and Equation (4.10) can be integrated to give

$$\log\left(\frac{N}{N_0}\right)=\frac{D}{I}(t-t_0) \tag{4.11}$$

where N_0 is the number of revolutions per second at time t_0 and N the number at a time t. The damping force can be computed by considering the tangential velocities of molecules hitting the ball surface. D is found to be linearly dependent on P as might be expected from Equation (4.7). Beams *et al.*[40] give an approximate formula

$$\log\left(\frac{N}{N_0}\right)=\frac{5P}{r\rho}\left(\frac{M}{2\pi kT}\right)^{1/2}(t-t_0) \tag{4.12}$$

where ρ is the density of the ball and r its radius. A similar formula has been derived by Fremerey[42].

Since at pressures of the order of 10^{-3} Pa it takes an hour or more for the ball to lose 1% of its velocity, the gauge requires an accurate measure of velocity. Also, precautions have to be taken to ensure thermal equilibrium and extraneous external fields must be eliminated or compensated; the earth's magnetic field, for example, can give an error equivalent to a pressure reading

of 10^{-5} Pa. It is usually assumed that the exchange of tangential momentum of the gas molecules at the rotor surface is perfect. The validity of this assumption has been discussed by Fremerey[43]. It is concluded that the assumption is justified for normal smooth metal balls at reasonable speed. However, at higher speeds and with rough surfaces there could be departure from the assumption. In particular it was shown by Comsa *et al.*[44] that with a roughened surface there was a greater variation in the measurements with gas species.

Pressures down to 10^{-7} Pa have been measured, but the gauge is most useful over the 10^{-2} to 10^{-5} Pa range. Even with the rugged versions, the early viscosity gauges were very much precision instruments requiring relatively expensive control and measuring electronics and which had to be mounted on antivibration tables. The practicability of the spinning rotor gauge was significantly enhanced by the introduction of small permanent magnet suspensions, for example by Fremerey and Boden[45] (see *Figure 4.18*). Gauges of this type are available commercially on a limited scale.

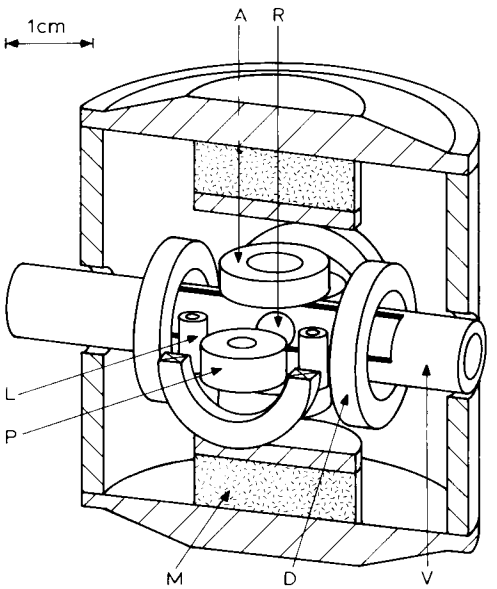


Figure 4.18 Permanent magnet ferromagnetic suspension spinning ball gauge head. R, Rotor; V, vacuum enclosure; M, one of two permanent magnets; A, one of two coils for pick-up and control of axial rotor position; L, one of four coils for lateral damping system; D, one of four drive coils; P, one of two pick-up coils

Radiometer gauges offer the advantage of being independent of the gas species. Their study, as the name implies, dates back to the original Crookes radiometer of 1873, a four-vaned device which rotated when a source of light or heat was brought near it and which is still to be seen in novelty shops today. However, it was not until 1914, when Knudsen[46] arrived at a clear explanation of the radiometer mechanism, that a gauge was developed based on the

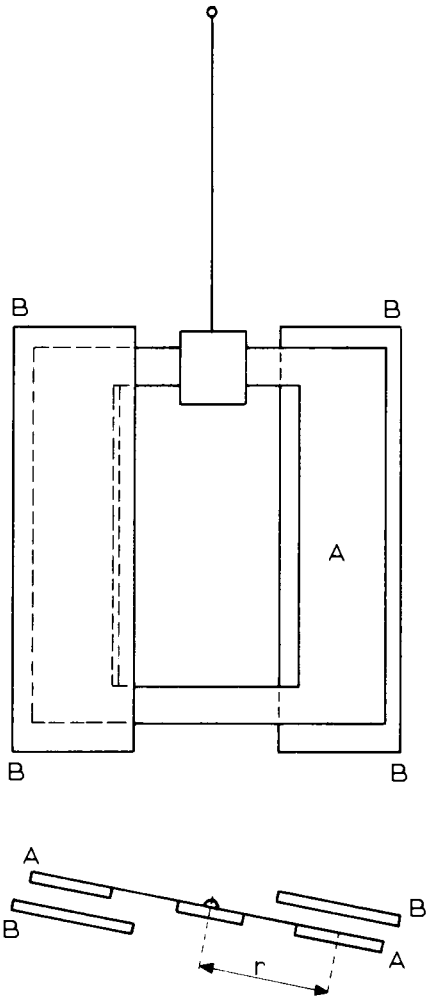


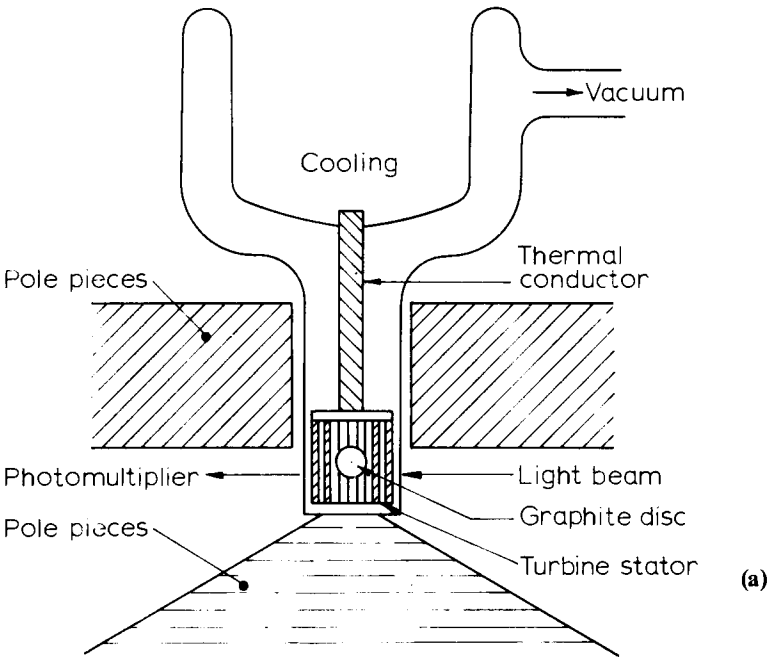
Figure 4.19 Schematic diagram of the Knudsen type of gauge

momentum transfer system between surfaces at different temperatures. The gauge of Knudsen consisted of a suspended vane placed close to two surfaces which could be heated, *Figure 4.19*. The deflection of the vane is proportional to the pressure, Equation (4.9) giving the form[39]

$$P = \frac{4\pi^2 I\delta}{rA\tau^2}\left(\frac{T_0}{T_1 - T_0}\right) \tag{4.13}$$

where I is the moment of inertia of the vane, A its area and r its mean radius. τ is the period of vibration of the vane and δ the deflection. T_1 is the temperature of the plates.

Several modified versions of the Knudsen gauge and other configurations exploiting the radiometer effect have been described over the years. Again the early designs are described in Dushman[39]. As with the viscosity gauges their main limitation was in the delicate suspension required for low pressures. However a radiometer gauge has been described by Evrard and Boutry[47]



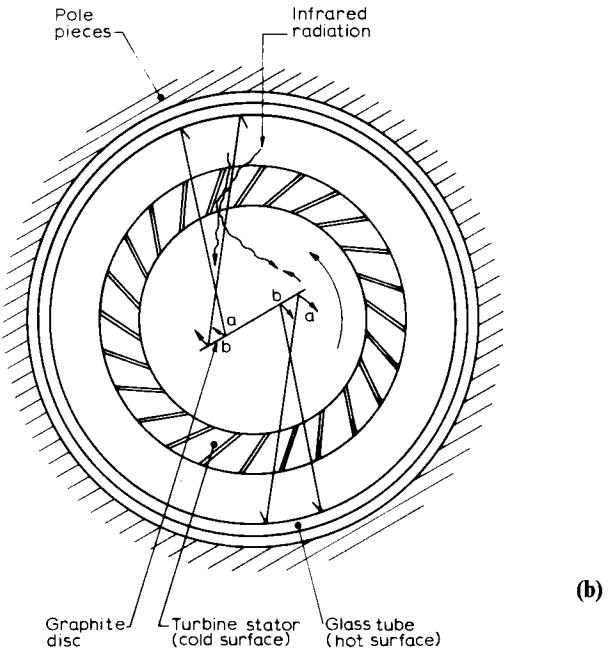


Figure 4.20 Schematic diagram of the levitation gauge of Evrard and Boutry[47]. (a) Plan view; (b) Horizontal cross section demonstrating the momentum transfer mechanism

which exploits the magnetic suspension principle and which is claimed to measure down to 10^{-8} Pa. In this gauge a disc of 3 mm diameter and thickness 0.05 mm is cut from polycrystalline, magnetically isotropic graphite and is set spinning about a diameter on a vertical axis. The disc rotates inside a cylindrical cage composed of molybdenum strips set at an angle as illustrated in *Figure 4.20*. The base of the cage is in good thermal contact with a metal rod which can be immersed in liquid nitrogen. Thus the cage will be at, say, 78 K whilst the envelope of the gauge is at ambient.

If the disc rotates in the direction of the arrow, *Figure 4.20(b)*, then the half-disc faces (a) are receding whilst the other half-disc faces (b) are advancing. Surfaces (a) are in direct line of sight of molecules which have collided with the relatively hot envelope surface as well as those from the cage, whereas molecules from the 'hot' surface cannot reach (b) without first colliding with the cage, i.e. all the molecules hitting (b) will have come from the cold cage-surface. Thus a torque proportional to pressure will be applied to the disc, due to the momentum transfer from the 'hot' molecules, which will accelerate or decelerate the rotating disc according to its initial direction of rotation. The rotation is measured by a light beam incident on the disc and the reflection measured on a photomultiplier. Two pulses of light are received for one rotation of the disc. The deceleration mode is normally used and the speed is determined by measuring the time interval of a number of pulses. At 10^{-7} Pa, a time interval of around 60 s is required, so that the gauge cannot be used where the pressure in the system is varying rapidly. By holding the cage at ambient temperature the gauge can be used as a viscosity gauge. In any case the viscosity effect has to be taken into account in the calculations.

Such a gauge is complex and expensive but, nevertheless, it represents a true standard gauge and has been used as such for calibrating ionization gauges.

4.4 Gauge calibration

There are basically two methods of calibrating gauges, a static method and a dynamic method. In the static method, the gauge to be calibrated is mounted beside a standard gauge in a vacuum chamber and the gauges compared under equilibrium conditions. The chamber must be designed to ensure equality of pressure in the region of the two gauges, i.e. no gas flow, and the system must be capable of maintaining a vacuum well below the pressures to be measured.

This method is particularly suitable for gauges measuring pressures in the 10 to 10^{-3} Pa range where sealed-off systems can be used. For such gauges accuracy of calibration of better than $\pm 1\%$ can be achieved. However, for ultrahigh vacuum gauges, fulfilling these conditions is almost an impossible task. First, the only gauges which can be used as standard are the momentum transfer gauges (Section 4.3) and they have a low-pressure limit of around 10^{-7} Pa. Thus, one has to rely on secondary standards, pre-calibrated ion gauges, or extrapolation for lower pressures. Secondly, the background pressure is required to be below 10^{-8} Pa, which is no mean task for any sealed-off vacuum system and it is difficult to ensure that the gas being used for calibration is indeed the known gas being introduced. Nevertheless, the

method is used and is satisfactory for calibration of gauges for many applications. In particular, it is useful for checking the sensitivity of production gauges. The problem of a suitable standard gauge can, to some extent, be overcome by using an expansion method. A small quantity of gas of accurately known volume and at a pressure above 10^{-3} Pa which can thus be measured more accurately, is expanded into a larger chamber containing the gauge to be calibrated and the pressure is calculated from the volume ratio. A volume ratio of up to 10^7 can be employed or, alternatively, a series of expansions can be used, employing several chambers, with smaller volume ratios. Such systems have been detailed by Holanda[48] and Elliot *et al.*[49].

The dynamic method is an extension of the expansion method. The two chambers are connected via a conductance of known value, C, and the gas is pumped with a measured throughput, Q Pa $m^3 s^{-1}$. If the pressure in the first chamber, used as the calibration chamber, is P_1 and in the second chamber P_2, then under equilibrium conditions

$$Q = C_1(P_1 - P_2) = SP_2 \tag{4.14}$$

where S is the pump speed at the second chamber and where molecular flow conditions through the conductance are assumed.

Various arrangements for embodying the method have been described in the literature, the differences hinging mainly on the method of measuring the parameters of Equation (4.14). Most use a circular orifice as the control conductance since its value can be calculated fairly accurately. The main problem with ultrahigh vacuum calibration is the conflict between obtaining sufficient pressure difference across the orifice with the need to achieve low base pressures. Clearly the throughput must be large compared with the outgassing rate of the system, but also a P_2 value of at least an order better than the lowest calibration pressure should be attainable. Poulter[50] describes a calibration system with a cryopump directly behind the calibration orifice, so that the molecules incident on the orifice had a very high probability of capture. By this method the base pressure, P_2, could be reduced to 2×10^{-9} Pa. Similar systems have been used by other workers. Another difficulty with gauge calibration at ultrahigh vacuum is the measurement of Q. At pressures around 10^{-4} Pa, the flow rate is of the order of 10^{-5} Pa m^3 and is measurable at the input fairly accurately using a flow meter. At 10^{-7} Pa the flow rate becomes, say, 10^{-8} Pa m^3 which is outside the range of simple flowmeter measurement. An alternative method to the flowmeter is to measure the pressure drop across a second orifice at the inlet whence

$$Q = C_0(P_0 - P_1) = C_1(P_1 - P_2) = SP_2 \tag{4.15}$$

In practice, C_0 is chosen to be small so that P_0 can be in the 10^{-1} Pa range and more easily and accurately measured. One form of conductance that facilitates this is the so-called porous plug used for calibrating leak detectors, etc., and which had been advocated for gauge calibration by Leck[51]. The plug can be ceramic or a sintered material such as silicon carbide. The porous material is, in effect, a large number of fine capillary tubes in parallel and free molecular flow conditions can be maintained up to atmospheric pressure.

Close *et al.*[52] describe a dynamic system for ultrahigh vacuum gauge calibration consisting of three chambers with a porous plug connecting the first two and an orifice between the second and third. The apparatus is shown schematically in *Figure 4.21*. The three chambers were stainless steel cylinders of approx. 3×10^{-3} m^3 volume and all the seals were metal bakable type, as were the valves. The conductance C_0 was a porous plug of silicon carbide sealed in a Pyrex tube and C_1 was a 0.3 mm hole drilled in a copper gasket between two flanges. The stainless steel tubes through which the gas flowed into chambers 2 and 3 were turned through a right angle to prevent sreaming and satisfy the requirements for a Maxwellian velocity distribution.

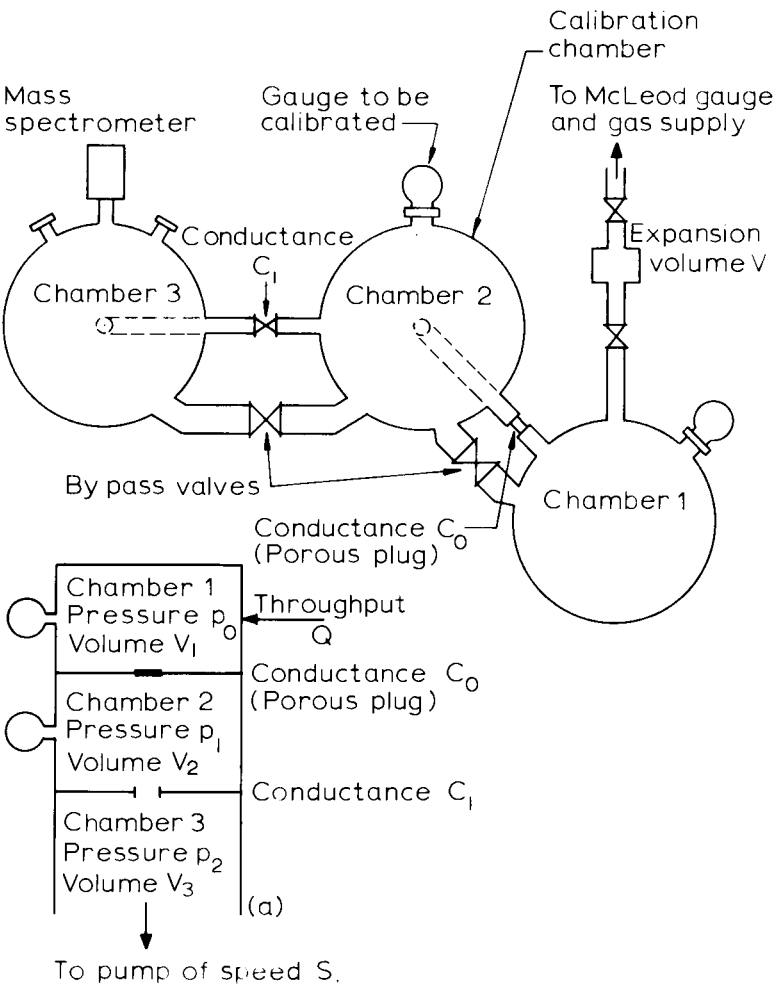


Figure 4.21 Schematic diagram of the ultrahigh vacuum gauge calibration system of Close *et al.*[52]

By arranging the pump speed to be much greater than C_1 and C_1 to be greater than C_0 (C_1/C_0 was about 10^5), Equation (4.15) reduces to

$$Q = C_0 P_0 = C_1 P_1 = S P_2 \tag{4.16}$$

and therefore

$$P_1 = \frac{C_0}{C_1} P_0 \tag{4.17}$$

Since the pressure depends on the ratio of C_1 to C_0, and both conductances vary in the same way with temperature and molecular mass of the gas, the system, once calibrated, is independent of the environmental temperature or the test gas being used. The initial pressure of the incoming gas was measured in the 10 Pa range with a McLeod gauge and a static expansion system was employed to reduce it to $\sim 10^{-1}$ Pa in the first chamber.

Important for calibration at ultrahigh vacua is a low outgassing rate and it is essential that the system and the gauges to be calibrated are bakable up to at least 250°C. Also it is important to run the ion gauges at low current to reduce the pumping effect. The accuracy of the various calibration methods and the sources of uncertainty are well described in a review article by Poulter[53]. For most applications using ultrahigh vacuum systems a gauge accuracy of $\pm 10\%$ is sufficient and with care, calibrations down to 10^{-8} Pa can be made with this accuracy. Below 10^{-8} Pa, extrapolation using modulator techniques and/or comparing several gauges, etc. must be resorted to, and as a result there is less certainty in the accuracy. It should be realized, however, that the sensitivity of a BAG can vary with time and that the gauges need to be calibrated regularly if the accuracy is to be maintained. Poulter and Sutton[54] found that the sensitivity of BAGs could vary by as much as 25% during their lifetime. This, to some extent, throws into doubt the usefulness of the BAG as a secondary standard for calibration purposes, and other designs have been suggested[55,56,57] which are claimed to be more stable. In general, such designs are not available for practical, ultrahigh vacuum use.

Ideally the gauge should be calibrated for all gases but in practice all that is required is calibration for one gas and knowledge of the relative sensitivities for the other gases. Although many attempts have been made to relate the relative sensitivities for the different gases to their physical properties and there does seem to be some correlation between relative sensitivity and the ionization cross section, there is no universal law which will allow accurate prediction of such values for all gauges. In *Table 4.1* relative sensitivities are given for several gauges taken from the literature. The sensitivities are expressed relative to nitrogen, this being the gas commonly used for calibration. The data indicate reasonable agreement for the BAGs in spite of differences in geometry and operating voltages but the other types of gauge show significant departures. Holanda[62] has collected data on four types of gauge and has looked for correlation with various gas properties. He found that ionization cross section correlated best with gauge sensitivity but that the optimum cross section value to choose (maximum value or the value at 2/3 accelerating potential for example) depended on the gauge type.

Bartmess and Georgiadis[64] measured the sensitivity of BAGs for 75 different gases covering a large number of organic compounds as well as inert and active gases. The ionization cross section was not known for many of the gases

TABLE 4.1. Sensitivities for various gauges relative to the sensitivity for nitrogen

	Gauge type and reference					
	Triode (Dushman and Young[58])	*BAG* (Schulz[59])	*BAG*** (Bennewitz and Dolmann[60])	*BAG* (Utterback and Griffith[61])	*BAG* (Holanda[62])	*Magnetron** (Barnes *et al.*[63])
He	0.158	0.21	(0.134)	0.180	0.18	0.24
Ne	0.24	0.33	(0.258)	0.31	0.32	–
A	1.19	1.50	(1.00)	1.42	1.42	1.76
N_2	1.00	1.00	–	1.00	1.00	1.00
H_2	0.46	0.42	(0.30)	0.423	0.41	0.52
O_2	–	–	–	0.874	0.78	0.99
CO	–	–	–	1.11	1.01	–
CO_2	–	–	(0.9)	1.43	1.39	1.29

* The value for the magnetron were obtained by calibrating against a BAG and using data from Leck[51]
** Sensitivities are relative to argon

but they found reasonable correlation between the relative sensitivity and polarizability which was known to be a linear function of the ionization cross section.

Calibration with nitrogen or the inert gases presents no great problem and reasonable accuracy can be attained. With the active gases such as oxygen and hydrogen, reaction with the gauge and chemisorption on the walls of the calibration system can introduce serious errors. Unfortunately, it is the active gases that can be important in an ultrahigh vacuum system and even if the relative sensitivities can be assumed, the same problems of chemisorption and gauge interaction will apply to the operation of the gauge when measuring pressure. The degree of confidence in pressure measurements when the residual gas is likely to be an active gas is therefore much less than for nitrogen.

In summary, we can say that the Bayard–Alpert gauge which has become established as a general-purpose ionization gauge covering pressures from 10^{-1} Pa down to 10^{-8} Pa can be calibrated over most of its range to an accuracy of better than $\pm 10\%$. The relative sensitivities for different gases can be deduced from data given in the literature with a reasonable degree of confidence. The incorporation of a modulator electrode allows the minimum pressure to be extended by an order or two, but with rather less degree of certainty. For most applications the accuracy is adequate and the main error may well be due to not knowing the composition of the residual gas. It is important in using such gauges to ensure that they are thoroughly outgassed and that the emission current used does not cause significant pumping.

Alternative ionization gauges such as the extractor gauge and the magnetron gauge are capable of measuring lower pressures but they are not readily available and are expensive. In general they are not suitable for higher pressures above 10^{-4} Pa which makes them less versatile and they are more difficult to calibrate.

The main drawbacks of the ionization gauge are that it is not an absolute gauge, the sensitivity depends on the gas composition and active gases can react with the gauge to cause errors. Momentum transfer gauges potentially can overcome these problems, but to measure in the ultrahigh vacuum region, precision configurations are required with sophisticated electronics. This

makes them unsuitable for general use and they have mainly found application as standards for calibration purposes.

Unless there is a big demand for better vacuum (below 10^{-10} Pa) in the future there is little need for further development of gauges for the ultrahigh vacuum region. Therefore, the only advances we are likely to see in the next few years will be minor modifications of existing designs aimed at improving performance, packaging or cost. Particularly we may expect to see modern electronic circuit techniques being applied to the control and measuring units, making them more compact and with digital output.

4.5 References

1. BAYARD, R. T. and ALPERT, D., *Rev. Sci. Instrum.*, **21**, 571, (1950)
2. NOTTINGHAM, W. B., *7th Annual MIT Conf. on Phys. Electronics*, (1947)
3. TATE, J.P. and SMITH, P. J., *Phys. Rev.*, **39**, 270, (1932)
4. LANDER, J. J., *Rev. Sci. Instrum.*, **21**, 672, (1950)
5. PITTAWAY, L. G., *J. Phys. D: Appl. Phys.*, **3**, 1113, (1970)
6. REDHEAD, P. A., *Vacuum*, **13**, 253, (1963)
7. VAN OOSTROM, A., *Trans. 8th Nat. Vac. Symp. & Proc. 2nd Internat. Cong. on Vac. Sci. Technol., 1961*, **1**, p. 443, Pergamon Press, (1961)
8. ALPERT, D., *J. Appl. Phys.*, **24**, 860, (1953)
9. CLOSE, K. J. and YARWOOD, J., *Vacuum*, **22**, 45, (1972)
10. HERBERT, B. K., *Vacuum*, **26**, 363, (1976)
11. LAFFERTY, J. M., *J. Appl. Phys.*, **22**, 299, (1951)
12. REDHEAD, P. A., *Rev. Sci. Instrum.*, **31**, 343, (1960)
13. HOBSON, J. P., *J. Vac. Sci. Technol.*, **1**, 1, (1964)
14. LANGE, W. J. and SINGLETON, J. H., *J. Vac. Sci. Technol.*, **3**, 319, (1966)
15. VAN OOSTROM, A., *J. Sci. Instrum.*, **44**, 927, (1967)
16. PITTAWAY, L. G., *J. Vac. Sci. Technol.*, **10**, 507, (1973).
    See also *A Novel Ion-Extraction Ultrahigh Vacuum Gauge*, Ph.D. Thesis, CNAA, (1981)
17. GROSZKOWSKI, J., PYTKOWSKI, S. and TRZOCK, W., *Proc. 7th Internat. Vac. Cong.*, 121, (1977)
18. MIZUNO, H. and HORIKOSHI, G., *Proc. 7th Internat. Vac. Cong.*, 129, (1977)
19. SCHUEMANN, W. C., *Rev. Sci. Instrum.*, **34**, 700, (1963)
20. METSON, G. H., *Brit. J. Appl. Phys.*, **2**, 46, (1951)
21. REDHEAD, P. A., *J. Vac. Sci. Technol.*, **3**, 173, (1966)
22. PITTAWAY, L. G., *Philips Res. Rep.*, **29**, 261, (1974)
23. GROSZKOWSKI, J., *Bull. Acad. Polon. Sci.*, **14**, 1023, (1966)
24. GROSZKOWSKI, J., *Le Vide*, **136**, 240, (1968)
25. PITTAWAY, L. G., *Philips Res. Rep.*, **29**, 283, (1974)
26. WATANABE, F., HIRAMATSU, S. and ISHIMARU, H., *Vacuum*, **33**, 271, (1983)
27. HELMER, J. C. and HAYWARD, W. H., *Rev. Sci. Instrum.*, **37**, 1652, (1966)
28. PENNING, F. M., *Philips Tech. Rev.*, **2**, 201, (1937)
29. PENNING, F. M. and NIENHUIS, K., *Philips Tech. Rev.*, **11**, 116, (1949)
30. HOBSON, J. P. and REDHEAD, P. A., *Can. J. Phys.*, **36**, 271, (1958)
31. REDHEAD, P. A., *Can. J. Phys.*, **37**, 1260, (1959)
32. KAGEYAMA, K., MONMA, S., FUKASAWA, F. and KURIYAMA, N., *Proc. 6th Internat. Vac. Cong. Jap. J. Appl. Phys. Suppl. 2*, 109, (1974)
33. YOUNG, J. R. and HESSION, F. P., *Trans. 10th Nat. Vac. Symp. 1963*, 234, Macmillan, (1964)
34. NICHIPOROVICH, G. A. and KHANINA, I. F., *Proc. 4th Internat. Vac. Cong. 1968*, 666, Inst. of Phys., (1968)
35. LAFFERTY, J. M., *J. Appl. Phys.*, **32**, 424, (1961)
36. LAFFERTY, J. M., *Proc. 4th Internat. Vac. Cong.*, 647, Inst. of Phys., (1968)
37. MEYER, E. A. and HERB, R. G., *J. Vac. Sci. Technol.*, **4**, 63, (1967)
38. WALKER, R. and PACEY, D. J., *Proc. 4th Internat. Vac. Cong. 1968*, 638, Inst. of Phys., (1968)
39. DUSHMAN, S., *Scientific Foundations of Vacuum Technique*, Wiley, (1962)

40. BEAMS, J. W., YOUNG, J. L. and MOORE, J. W., *J. Appl. Phys.*, **17**, 886, (1946)
41. BEAMS, J. W., SPITZER, D. M. and WADE, J. P., *Rev. Sci. Instrum.*, **33**, 151, (1962)
42. FREMEREY, J. K., *J. Vac. Sci. Technol.*, **9**, 108, (1972)
43. FREMERY, J. K., *Vacuum*, **32**, 685, (1982)
44. COMSA, G., FREMEREY, J. K. and LINDENAU, B., *Proc. 7th Internat. Vac. Cong.*, 157, (1977)
45. FREMEREY, J. K. and BODEN, K., *J. Phys. E. Sci. Instrum.*, **11**, 106, (1978)
46. KNUDSEN, M., *Ann. Phys.*, **44**, 525, (1914)
47. EVRARD, R. and BOUTRY, G. A., *J. Vac. Sci. Technol.*, **6**, 279, (1969)
48. HOLANDA, R., *NASA Tech. Note No. NASA TND 3100: N69-35265*, (1969)
49. ELLIOTT, K. W. T., WOODMAN, D. M. and DADSON, R., *Vacuum*, **17**, 439, (1967)
50. POULTER, K. F., *J. Phys. E: Sci. Instrum.*, **7**, 39, (1974)
51. LECK, J. H., *Pressure Measurement in Vacuum Systems*, 132, Chapman & Hall, (1964)
52. CLOSE, K. L., VAUGHAN-WATKINS, R. S. and YARWOOD, J., *Vacuum*, **27**, 511, (1977)
53. POULTER, K. F., *J. Phys. E: Sci. Instrum.*, **10**, 112, (1977)
54. POULTER, K. F. and SUTTON, C. M., *Vacuum*, **31**, 147, (1981)
55. BILLS, D. G., ARNOLD, P. C., DODGEN, S. L. and VAN CLEVE, C. B., *J. Vac. Sci. Technol.*, **A2**, 163, (1984)
56. CHOUMOFF, P. and IAPTEFF, B., *Proc. 6th Internat. Vac. Cong. 1974, Japan J. Appl. Phys. Suppl. 2, Part 1*, 143, (1977)
57. SUTTON, C. M. and POULTER, K. F., *Vacuum*, **32**, 247, (1982)
58. DUSHMAN, S. and YOUNG, A. H., *Phys. Rev.*, **68**, 278, (1945)
59. SCHULZ, G. J., *J. Appl. Phys.*, **28**, 1149, (1957)
60. BENNEWITZ, H. G. and DOLMANN, H. D., *Vakuum Techn.*, **14**, 8, (1965)
61. UTTERBACK, N. G. and GRIFFITH, T., *Rev. Sci. Instrum.*, **37**, 866, (1966)
62. HOLANDA, R., *J. Vac. Sci. Technol.*, **10**, 1133, (1973)
63. BARNES, G., GAINES, J. and KEES, J., *Vacuum*, **12**, 141, (1962)
64. BARTMESS, J. E. and GEORGIADIS, R. M., *Vacuum*, **33**, 149, (1983)

Chapter 5

Partial pressure measurements

5.1 The need for partial pressure gauges

For many applications it is sufficient to know the order of the residual pressure in terms of the equivalent nitrogen pressure and for this a total pressure gauge can be used without the need for constant calibration. If, however, the pressure is required to the accuracy of the ion gauge then not only must the gauge be regularly calibrated and the relative sensitivity for different gases be known but the composition of the residual gas in the system must also be ascertained.

At pressures above 10^{-5} Pa, especially when pumping with a diffusion pump–rotary pump combination, the residual gas is likely to have a similar composition to the atmosphere present before evacuation. In applications in this pressure range the pressure can be measured reasonably accurately and the total pressure values are meaningful. Below 10^{-5} Pa this may no longer be the case.

It was pointed out in Chapter 3 that some of the pumps used for ultrahigh vacuum systems are selective when pumping gas mixtures, often leaving some gases virtually unpumped. Also, at pressures below 10^{-6} Pa the gas evolved from the components of the vacuum system plays a major role in determining the ultimate pressure and this gas may bear little or no relation to the original gas atmosphere. Further, for many applications, the nature of the residual gas can be more important than its actual pressure; active gases are a problem in surface studies, whilst the presence of heavy molecules can be detrimental in particle accelerators. Under these conditions, measurement of the total pressure may not be enough to give meaningful information about the vacuum environment. There is a need to know at least what the main constituent gas is and ideally what other gases are present and their relative abundance or, better still, their partial pressures. As a result, in parallel to the development of total pressure gauges for ultrahigh vacuum, there has also been development of partial pressure gauges to operate over the same pressure range.

A partial pressure gauge is basically a mass spectrometer designed for investigating residual gases in a vacuum system. Sometimes termed a residual gas analyser, it may or may not be calibrated for direct partial pressure measurement. It differs from more conventional mass spectrometers in having

a higher sensitivity and in being more compact. On the other hand, the mass range and resolution are relatively limited.

Essentially, ions are formed in a source by electron impact ionization of the residual gas, as in an ion gauge. The ions are extracted with suitable fields and passed through an analysing section which separates the ion species before they are collected and measured. Normally a single collector is employed and some parameter of the analyser, such as the electric field or magnetic field, will be varied to 'scan' the mass range across it. The output of the collector is then plotted to give a 'spectrum' such as shown in *Figure 5.1*. In this way each ion

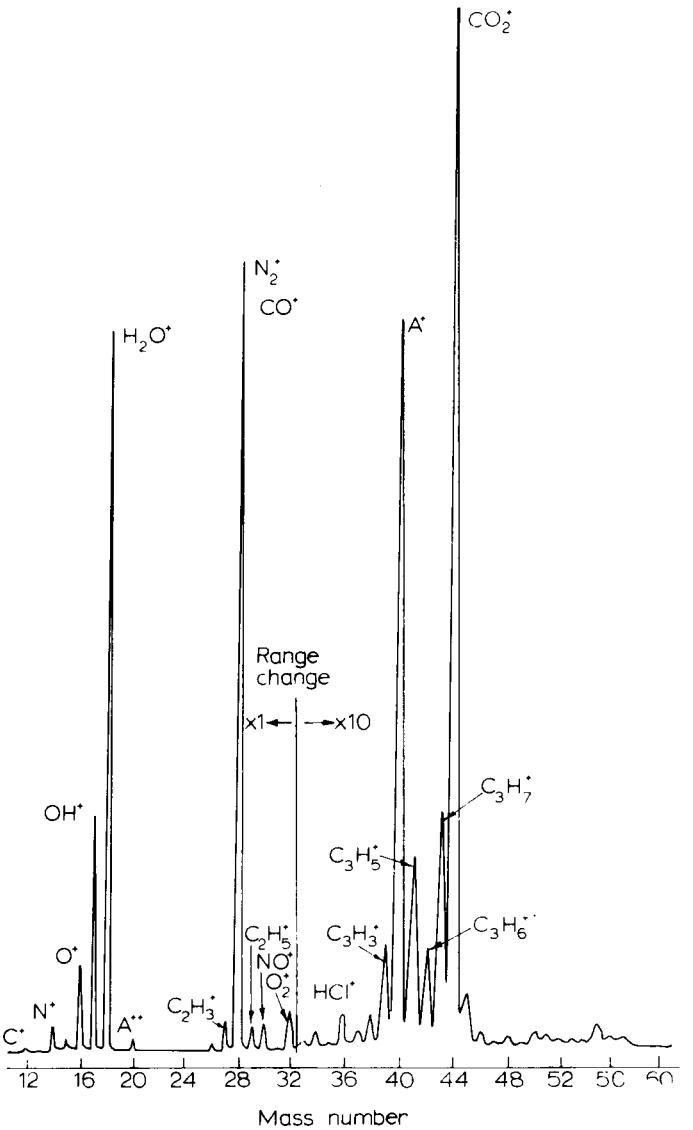


Figure 5.1 Typical residual gas spectrum measured with a 180° magnetic sector spectrometer

species can be selected and by a suitable calibration of the collector current the partial pressure can be determined. In other words, the peak height × the sensitivity for a specific gas should be a measure of the partial pressure.

In practice it is not quite as simple as that. In the ionization region electron impact produces not only singly charged ions but also doubly charged ions and dissociated fragment ions. Separation of the ions in the analyser depends on the charge-to-mass ratio and, therefore, each polyatomic gas present will produce an output at several values of charge-to-mass ratio. This can be seen in *Figure 5.1* for composite gases such as CO_2 and water vapour and for diatomic molecules such as N_2. Also, ions of different gases can have the same charge-to-mass ratio, e.g. N_2^+ and CO^+. Thus, although the determination of the gases present is fairly readily ascertained from such spectra, their relative abundance and, hence, the actual partial pressure of each gas is more difficult to determine with any great degree of confidence. This is discussed further in Section 5.7 where calibration is considered. Fortunately, for most applications accurate partial pressure measurement is not essential and combining the information from a residual gas analyser with total pressure measurement is adequate.

5.2 Gauge parameters

The performance of a partial pressure gauge will clearly depend on criteria such as the ionization and ion extraction efficiencies, the degree of separation of the various gas ions, and the detection efficiency. These criteria can be characterized by the following gauge parameters, which normally, although not always, would be quoted by the designer or manufacturer:

(1) Mass range.
(2) Pressure range.
(3) Mass resolution.
(4) Sensitivity.
(5) Minimum detectable partial pressure.
(6) Partial pressure sensitivity.

The mass and pressure ranges are fairly straightforward and are usually given as absolute limits without regard to resolution or sensitivity. The resolution is less clearly defined as the ability to separate adjacent mass peaks. It is expressed as the ratio of $M/\Delta M$, where M is the ion mass number equal to m/m_0 where m is the ion mass and m_0 the mass of a reference ion and ΔM the minimum mass difference between two peaks that can still be resolved. Thus, we can expect a variation in resolution with mass number and a dependence to some extent on relative peak heights. In practice, the definition is simplified by taking ΔM equal to 1 atomic mass unit (amu) and we then speak of unit mass resolution when two peaks of mass M and $M+1$ are resolved.

What is meant by resolved is, to some extent, controversial, and various criteria are used. One common approach is to consider two adjacent peaks M and $M+1$ of equal height and measure the height of the valley between them as a percentage of the peak height, *Figure 5.2(a)*. For accurate measurements, especially when adjacent peaks are of unequal height, a 1 or 2% valley height is required whilst for semi-quantitative results a 50% valley is acceptable. A 10% valley height is a reasonable compromise and is most often used. Thus the

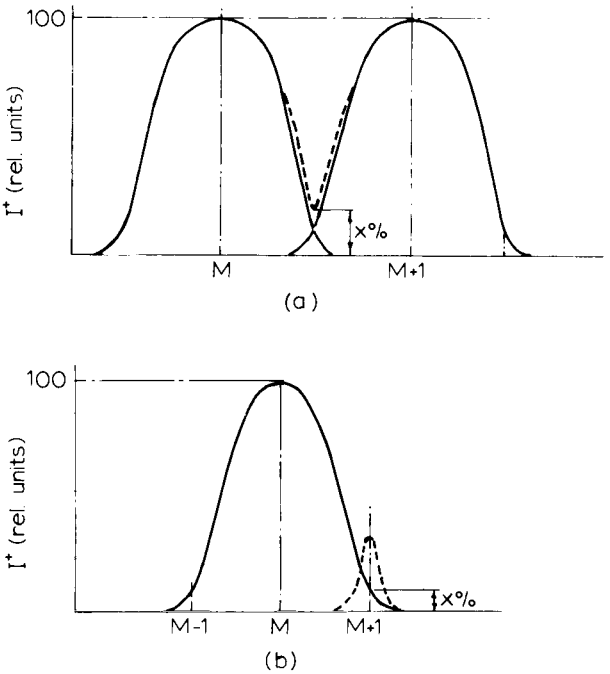


Figure 5.2 Definition of resolving power. (a) $x\%$ valley definition; (b) Resolution with $x\%$ contribution to adjacent mass

resolution of a residual gas analyser may be quoted as 'unit mass resolution at 10% valley' over the mass range of 2–200 amu or simply as 10% valley resolution over the 2–200 amu range. Sometimes it will just be quoted as resolution or resolving power (10% valley) = 200, since for most spectrometers the highest mass that can be resolved is the limiting factor.

An alternative approach is to measure the percentage height of a peak of ion mass M contributing to adjacent peaks at mass $M+1$ or $M-1$, *Figure 5.2(b)*. Both definitions can be assessed by examining the peak profile of a single ion mass M, and measuring the width at the appropriate percentage of the peak height.

The sensitivity of a mass spectrometer is usually defined as the ratio of the ion current to the gas pressure in the source. This definition is not in accordance with the term used for ionization gauges, since it is a function of the ionizing electron current. However, in most spectrometers the conditions of the source are optimized and set, so that the definition is fairly meaningful. It should be noted that the sensitivity will vary for different mass peaks due to ionization efficiency, etc., and therefore the gas must be specified for the sensitivity. Normally the sensitivity for nitrogen would be quoted. Relative to an ion gauge the sensitivity of most spectrometers is low. This results from the small source size often needed to give a low energy spread to extracted ions. However, since the collector is well screened from the ion source there is no problem of X-ray current and the sensitivity can be enhanced by incorporating an electron multiplier (see Section 5.6). The minimum current that can be

detected dictates the minimum partial pressure P_{min} that can be determined. The partial pressure sensitivity is then defined as the ratio P_{min} to the total pressure in which it can still be detected.

Having defined the parameters, the requirements for an ultrahigh vacuum partial pressure gauge can be considered. As with the ionization gauge, the instrument should not act as a source of gas or as a pump. It must, therefore, be bakable up to 200°C at least, without changing its characteristics. If the electronics or magnets have to be removed for baking, removal and replacement should not require recalibration. The gauge should have a minimum degassing rate and not contain components with large surface areas. Also the components should be made of materials compatible with ultrahigh vacuum practice, with the envelope having negligible gas permeability. The emission current should be low to prevent excessive pump action and heating. For most applications the gauge will be used for examining changing phenomena such as gas analysis during pumping. It should, therefore, be capable of reasonably fast scanning. This is not just reflected in the scanning technique and current measuring circuitry, but also in the volume capacity of the spectrometer. The source must be kept as small as possible with good gas conductance between source, analyser and detector. The sensitivity must allow ultrahigh vacuum pressures to be measured. On the other hand, the mass range is rarely required to be greater than 1–50 amu with 10% valley resolution. Although gases with mass numbers above 50 may be required for some applications, a resolution of 50% valley in this region is often adequate. The gauge should be easy to operate without complex electronic controls. Preferably the gauge heads should be interchangeable so that one control unit can be used with more than one head. Finally, the price should be economic enough to allow the gauge to be used as a general purpose instrument for ultrahigh vacuum systems to complement total pressure gauge measurements.

The development of mass spectrometers for chemical analysis and isotope separation dates back to the experiments of J. J. Thomson and Aston at the turn of the century. However, exploitation of spectrometers for vacuum environmental studies did not occur until the 1950s. At that time commercial spectrometers were fairly large sophisticated magnetic-sector type instruments, continuously pumped with residual gas pressures in the region of 10^{-5} Pa. They certainly did not meet the requirements for residual gas analysis outlined above. The need to know the gas environment in the extensive high vacuum investigations in the 1950s prompted the design and construction of much smaller versions of the traditional mass spectrometers, sacrificing resolution to achieve the above requirements on sensitivity, background pressure and scanning rate. As a result there are today a variety of magnetic-sector type partial pressure gauges available commercially, from the cheap instruments useful for indicating the residual gas composition, to the more accurate instruments for partial pressure measurements with higher resolution. During the same period there was also considerable investigation of alternative mass separation techniques some of which had been made feasible by the development of electronic circuitry. As example, 'time-of-flight' instruments gained popularity as electronic timing techniques became available. From these studies several new types of analyser were developed and are now being exploited. The types of analyser at present available can be divided into two groups—static spectrometers and dynamic spectrometers.

In the former, ions are separated by magnetic and/or electric fields which remain constant with time and will only be varied to scan the spectrum. The group covers magnetic and electrostatic sector instruments and the cycloidal spectrometer. In the dynamic spectrometers, time dependence of one or more of the system's parameters is essential for separation of the ions. This group includes r.f. spectrometers, time-of-flight instruments, the quadrupole and the omegatron.

Although many of these designs are available today and will be described in this chapter, two main types have emerged as general-purpose instruments, namely the magnetic sector type and the quadrupole. More emphasis will, therefore, be given to these two instruments.

5.3 Ion sources

Most gauges require a similar ion source and it is sensible to discuss this as a separate issue. The commonest ion source uses a beam of electrons from a thermionic emitter to bombard and ionize the residual gases present and includes a means of extracting and focusing the ions. Since many of the instruments are single focusing, i.e. they will focus ions with different path directions but not ions of different energies, an ion source is required with a small energy spread to give good mass resolution. This implies that the ionization must take place in a region over which the potential variation is very small and that the ions must be extracted without further deterioration of the energy spread.

A number of designs have been suggested in the literature but it is interesting to note that nearly all mass spectrometers of the magnetic sector type and some other types as well, use ion sources based on a design of Nier[1] introduced in 1940. The general principle of a Nier-type source is illustrated in *Figure 5.3*. The electrons are emitted from a filament lined up with slits in the accelerating plates A and B, the slits being perpendicular to the direction of the ion beam. These plates are held positive with respect to the filament and control the emission and energy of the electrons. The electron beam in the form of a 'ribbon' is prevented from diverging by a parallel magnetic field. In some designs use is made of the analysing magnetic field, but in others a separate field is provided by small permanent magnets incorporated in the source. The beam passes through the ionizing section or chamber, then through an exit slit in plate C, to be collected at the anode trap D. Ions formed in the electron beam are drawn out by a small field of the order of a few volts per centimetre between the plate R (repeller) and E. They are then accelerated by a high field between E and F of up to 1 kV to give the ions a common kinetic energy for the analysing system which is large compared with their initial energy spread. In certain designs a ribbon beam is required, in which case the defining slits in E and F will be rectangular with the long axis parallel to the direction of the electron beam.

It is important that the ion source can be thoroughly outgassed as, apart from upsetting the partial pressure determination, electron bombardment will cause desorption of the gas as ions from the electrodes. The desorption ions will have a higher energy than the gas-phase ions formed by electron ionization (cf. Section 4.2.2) and will present a secondary peak of apparently

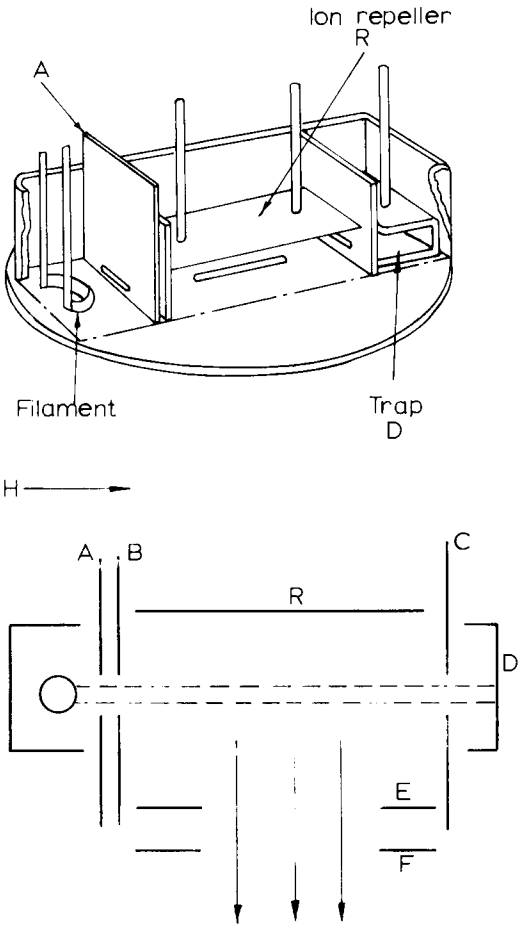


Figure 5.3 Typical Nier type ion source

different mass on the spectrum. It is also essential in the constructional design of the source to provide for replacement of the filament on 'burn out'. Normally a tungsten filament is used but also rhenium or lanthanum hexaboride coated rhenium can be used to minimize chemical reaction with the hot filament and also the heating of surrounding electrodes. The design and performance of such sources are well documented in the literature, for example in the book of Barnard[2], to which the reader is referred for more detailed information.

Although the electrons take a spiral path in the magnetic field, the sensitivity is low compared to an ion gauge, typically $S = 10^{-6}$ A Pa^{-1} which, since the electron current is restricted to about 200 μA, is down by a factor of four on that of a Bayard–Alpert gauge. This is only partly due to the lower ionization efficiency, since the number of ions formed that can be extracted (extraction efficiency) is also a limitation. The extraction efficiency depends on the aperture but, unfortunately, increasing the size of the aperture in general decreases resolution by increasing the energy spread of the ions.

An improvement in sensitivity has been suggested by Pittaway[3] by exploiting the extractor gauge. He modified his own design (*Figure 4.11*) by replacing the extractor electrode with two mesh electrodes, mounted perpendicular to the ion beam direction, to act as extractor and accelerator electrodes. He obtained a sensitivity of about 10^{-4} A Pa^{-1} with an electron current of 10 mA, but since the mesh was only 40% optically transparent a portion of the current was collected by the mesh wires. Pittaway indicated that a further improvement could be attained by employing a three electrode lens system for extracting and focusing the ions instead of the mesh electrodes, as illustrated in *Figure 5.4*.

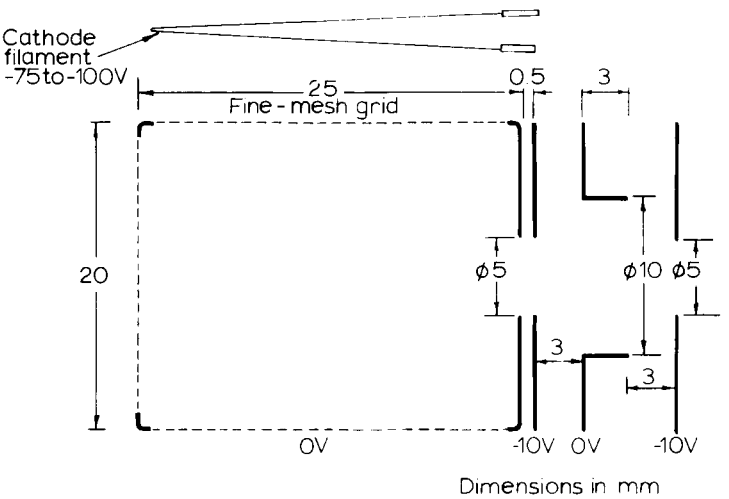


Figure 5.4 Schematic of an extractor-gauge type ion source with a three electrode electrostatic lens for ion extraction and focusing according to Pittaway[3]

For some analysers the energy spread is less critical and simpler ion sources will suffice; in particular there is no need for a magnetic field to focus the electrons. Whatever type of source is used, it is necessary to stabilize the emission current fairly carefully and control circuits similar to those used with ion gauges are required.

Having produced the ions with a small energy spread, the mechanism of separating the ion species can be considered. There are several methods of separating the ion masses most of which are unrelated. They are discussed under the headings of the spectrometer types, since the method employed in the analysing section defines the whole instrument.

5.4 Static spectrometers

5.4.1 Magnetic sector analyser

The magnetic sector analyser dates back to the early instrument developed by Dempster[4] in 1918 and represents the most important method of separating the ion species after they have emerged from the ion source.

If an ion is accelerated by a voltage V it will gain a velocity v given by $\frac{1}{2}mv^2 = eV$. If such an ion is launched into an homogeneous magnetic field H, at right angles to its direction of motion, it will experience a force of Hev, which is mutually perpendicular to the magnetic field and direction of travel of the ion. This force will cause the ion to follow a circular trajectory, the radius r of which can be determined by equation the electromagnetic force to the centrifugal force acting on the ion.

$$\frac{mv^2}{r} = Hev \tag{5.1}$$

or in terms of voltage V,

$$r = \frac{1}{H}\left(\frac{2mV}{e}\right)^{1/2} \tag{5.2}$$

Thus the radius of the ion trajectory depends on the mass-to-charge ratio of the ion m/e, the magnetic field and the accelerating voltage. If r is expressed in cm, H in gauss, V in volts and m is converted to M and expressed in amu for a singly charged ion,

$$r = 143.95(MV)^{1/2}/H \tag{5.3}$$

In Dempster's design the ions were deflected through 180° before collection in the arrangement illustrated in *Figure 5.5*. The magnetic field extends over the whole instrument and thus can also be used for focusing the electron beam in the source. The spectrum can be scanned by varying the accelerating voltage V or the magnetic field H, if an electromagnet is used. The magnetic field also has the effect of focusing a divergent beam and this is illustrated in *Figure 5.6* where trajectory plots are shown for ions with different initial directions. The focus is not perfect and the beam width at the exit from a mono-energetic line source is given by $r\alpha^2$, where α is the diverging angle from normal incidence of the input beam, for small values of α.

The energy spread will also increase the beam width. If the entry slit is of width S_1 then we can express the beam width at the exit as $S_1 + \delta S_1$ where δS_1 is the error due to the above aberrations. To collect all the ions of a given m/e, the exit slit width S_2 should be equal to or greater than $S_1 + \delta S_1$. For the

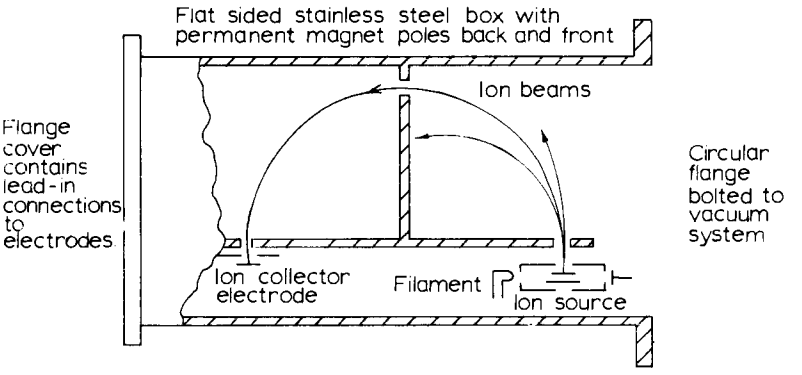


Figure 5.5 Schematic of 180° magnetic sector residual gas analyser

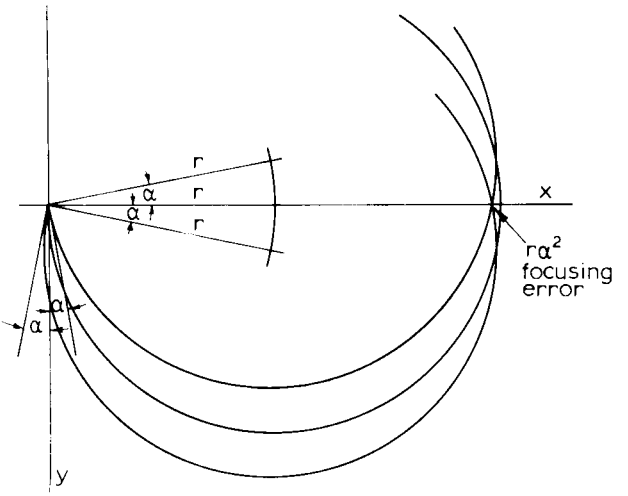


Figure 5.6 Focusing effect of the 180° sector showing the aberration $r\alpha^2$ for first-order focusing

greatest sensitivity, S_1 should be as large as possible. However, this would mean a large value of S_2 and overlapping of the output from ions of different masses. It can be shown, in fact, that the resolution is given by

$$\frac{M}{\Delta M}=\frac{r}{S_1+S_2+\delta S_1} \quad \text{or} \quad \frac{r}{2(S_1+\delta S_1)} \tag{5.4}$$

when $S_2=S_1+\delta S_1$. Thus the resolution varies directly as the radius and inversely as the slit width, and a compromise between resolution and sensitivity must be reached on the slit widths.

To use wide slits with good resolution, the early instruments employed a large beam deflection radius. However, for a compact residual gas analyser a much smaller radius is required. This not only restricts the slit widths but it also requires higher magnetic fields or lower ion accelerating voltages (cf. Equation (5.3)). The accelerating voltage needed to focus the ions on the exit slit decreases with increasing mass of the ion. However, since a lower limit is set by the thermal velocities of the ions, which must be small compared with the accelerating velocity, the mass range is limited by these requirements.

Nevertheless, 180° magnetic-sector analysers are commercially available with an ion deflection radius of only 10 mm, which give a range of 0–200 amu with 10% valley resolution up to mass 44. The sensitivity of such instruments is around 6×10^{-7} A Pa^{-1} allowing measurements in the ultrahigh vacuum region, down to partial pressures of 10^{-9} Pa. They are bakable up to 400°C and are relatively inexpensive. The scan rate, achieved by varying the acceleration voltage over two orders of magnitude, is of the order of 1 amu per second. The spectrum of *Figure 5.1* was taken with a small 180° magnetic spectrometer and is fairly typical of their performance.

Equation (5.3) applies equally to any magnetic field at right-angles to the ion trajectory, but there will not necessarily be a focusing effect as for the 180° sector. However, in the 1930s a closer look was made of the focusing criteria in

magnetic fields giving other than 180° deflection and Barber[5] and Stephens[6] demonstrated at about the same time, that a sector magnetic field giving the ions a deflection of less than 180° could have a focusing effect and be exploited for mass separation. The best focus is given when the incident beam and emergent beam are normal to the field boundaries and where the position of the entrance and exit slits are equal distance from the intersection of the two field boundaries and on a line through it.

This is illustrated in *Figure 5.7*. This is an idealized model with limited beam divergence and sharp boundaries to the magnetic field. Nevertheless, from the

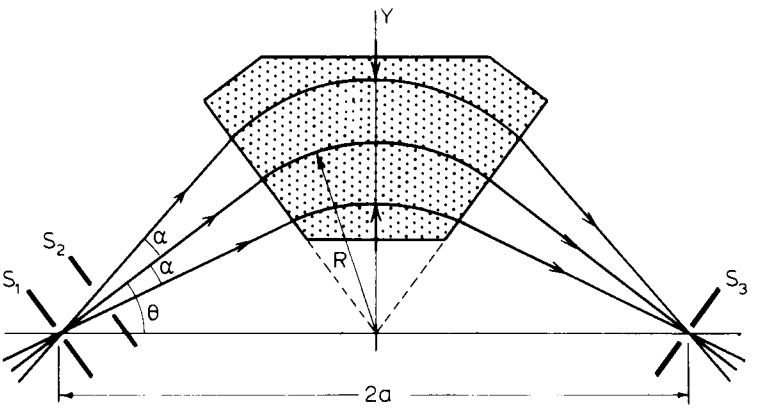


Figure 5.7 Magnetic sector analyser showing the position of source and collector for first-order focusing

work of Barber and Stephens and other investigations, more fully described in Barnard's book[2], mass spectrometers were developed from the 1930s with beam deflections of 30°, 60° and 90°. The advantage of the smaller angle sector type over the 180° sector spectrometer is that the collector and source can be outside the magnetic field. This facilitates the use of an electron multiplier to enhance the sensitivity of the collector and the source can be inserted into the vacuum system, in a similar way to a nude gauge.

The sector field can be obtained by pole pieces shaped to the required sector. The idealized situation of sharp magnetic boundaries cannot be achieved in practice and the fringing fields have to be taken into account. However, by suitably shaping the pole pieces, second order focusing can be obtained and this has been investigated both theoretically and experimentally[2]. As a result, the resolution and sensitivity of a 90° sector instrument, for example, can be comparable to or better than a 180° spectrometer of similar overall size. On the other hand, the alignment of the magnetic field is more critical and removal of the magnet for baking, etc. can upset the calibration.

As with the 180° instrument, the resolution depends on slit widths and the deflection radius. For high resolution, 60° magnetic-sector spectrometers are used with a large radius of curvature. For the residual gas analyser the 90° sector is usually employed. In 1960 Davis and Vanderslice[7] described such an instrument with a 5 cm radius for ultrahigh vacuum application. The instrument had a built-in electron multiplier and could measure partial

pressures down to 10^{-11} Pa. The input and output sections were mounted in glass envelopes which were attached to the metal section passing through the analysing magnet.

All metal (stainless steel) 90° sector spectrometers of similar size to that of Davis and Vanderslice (6.2 cm radius) are commercially available today which, by using an electromagnet, can cover a mass range of 500 amu with unit mass resolution up to 400 and with a sensitivity of 10^{-6} A Pa^{-1} for nitrogen. They present a more versatile and 'better performance' instrument than the cheaper small 180° magnetic sector instrument mentioned above and are particularly useful in vacuum systems where organic vapours are likely to be present.

Figure 5.8 shows photographs of two presently available magnetic sector type instruments.

5.4.2 Electrostatic sector analyser

For some applications the stray magnetic field from the magnetic sector analyser cannot be tolerated. Werner[8] has pointed out that a similar instrument using electrostatic deflection offers a possible alternative which has its attractions. In such an instrument the deflection is imposed on the ion by an electrostatic field E between two parallel plates, as shown schematically in *Figure 5.9*. As with the magnetic sector the radius of the circular orbit can be determined by equating the electric force on the ion against the centrifugal force

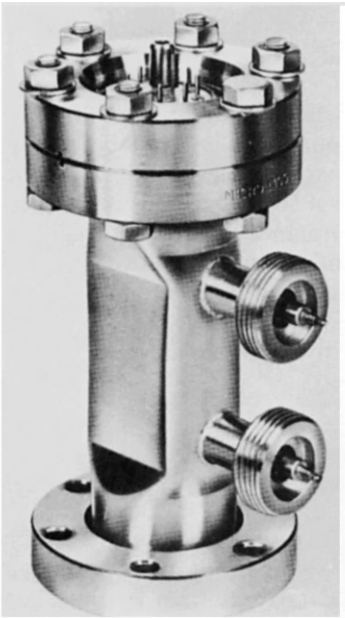
$$\frac{mv^2}{r} = eE \tag{5.5}$$

In the ion sources described in Section 5.3 all the ions escape with approximately the same kinetic energy equal to eV, where V is the accelerating voltage on the final extraction electrode (*Figure 5.3*). Thus, from Equation (5.5), it is evident that the ions from such a source will not be separated by the electrostatic deflection field but all will follow the same orbiting circle in the field. However, if the ion source is arranged to emit molecules with the same momentum rather than kinetic energy, then mv will be constant and from Equation (5.5)

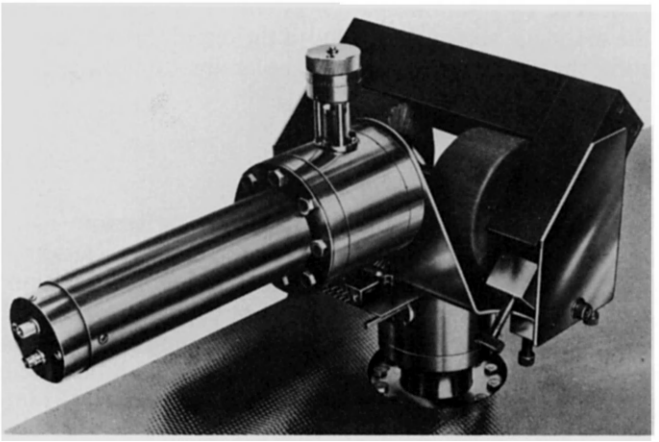
$$r \propto \frac{1}{meE} \tag{5.6}$$

Thus, molecules with different masses can be separated.

Essential, therefore, to the functioning of an electrostatic sector analyser is an ion source giving the gas molecules a constant momentum. Such a source is illustrated in *Figure 5.9*. The ions are generated by electron bombardment in the area between the electrodes R and L, placed distance d apart. A square-wave voltage is applied to L, negative relative to R, of duration τ and peak voltage V. τ is chosen to be small compared with the time of flight of the ions from R to L whilst the time between pulses must be greater than the flight time. Assuming the ions are singly ionized, the force applied to the ions will be eV/d and the ions will then all receive the same momentum



(a)



(b)

Figure 5.8 Two magnetic sector type spectrometers manufactured by VG Micromass Ltd. (a) Micromass 1, a 1 cm radius 180° sector instrument; (b) Micromass 6, a 6 cm 90° sector instrument

$$mv = \int_0^{\tau} \frac{eV}{d}\,\mathrm{d}t$$

$$= \frac{e}{d} V\tau \tag{5.7}$$

Combining Equations (5.5) and (5.7) gives the radius of the trajectory as

$$r = \frac{e}{m} \cdot \left(\frac{V\tau}{d}\right)^2 \cdot \frac{1}{E} \tag{5.8}$$

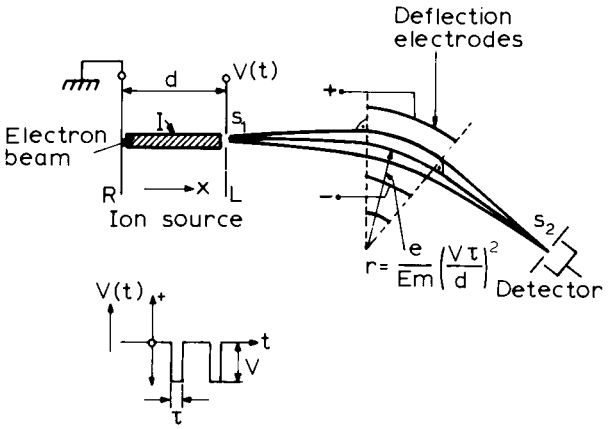


Figure 5.9 Principle of electrostatic sector analyser

The position of the slits S_1 and S_2 and also the values of V, E and τ are arranged so that only ions of a given mass will be selected and pass through slit S_2. By varying either V, τ or E the mass range can be scanned.

The fact that no magnets are required means that, apart from the absence of the stray magnetic field, the analyser head can be made lighter and more compact. Also, there is no longer the problem of having to realign the magnets after removal for baking. Werner points out other advantages. It is generally easier to produce ions with the same momentum than ions with the same energy and by changing the source field from a pulsed voltage to d.c. voltage on the accelerating electrode L, the analyser can be turned into a total pressure gauge. This is because the d.c. voltage produces ions with the same energy, which then follow the same path through the field. Of course the converse is also true, i.e. by applying a pulsed field to the magnetic sector instrument the ions will also all take the same path (Equation (5.2)). As far as the author is aware the idea of an electrostatic sector analyser has not been taken up commercially.

5.4.3 The cycloidal mass spectrometer

The cycloidal mass spectrometer is similar to the 180° magnetic-sector spectrometer but with an electric field applied at right-angles to the magnetic field.

The idea of using crossed magnetic and electric fields for a spectrometer was first described by Bleakney and Hipple[9] in 1938, and it was further investigated and improved in the 1950s for residual gas analysis.

The equation of motion for an ion launched into the crossed electric and magnetic field is given by

$$m\frac{d^2x}{dt^2} = He\frac{dy}{dt}$$
$$m\frac{d^2y}{dt^2} = Ee - He\frac{dx}{dt} \qquad (5.9)$$

If $y=0$ and $x=0$ at $t=0$, then Equation (5.9) can be integrated to give the solution

$$x = A\sin\theta - A\sin\left(\frac{eH}{m}t+\theta\right)+\frac{E}{H}t$$
$$y = A\cos\theta - A\cos\left(\frac{eH}{m}t+\theta\right) \tag{5.10}$$

where A and θ are constants of integration and depend on the initial velocity and direction of the ions relative to the electric field. Equation (5.10) is that of a cycloid, i.e. the path traced out by a point on a rolling wheel.

It will be noted from Equation (5.10) that at a time, $t=2\pi nm/eH$ where n is an integer, $y_n=0$ and $x_n=2\pi nmE/eH^2$, i.e. the position is independent of A and θ. Thus, ions of the same mass-to-charge ratio emerging from the point $x=0$, $y=0$ are all focused at a distance $x_n=2\pi nEm/eH^2$ from the origin regardless of the initial velocity or direction of the ions. It is thus the case of a perfect double focusing system and there is no longer a need for a source producing mono-energetic ions.

The special form of cycloidal path, known as the prolate form, has been mainly exploited for mass spectrometry. This is the case of the curve traced out by a point distance d from the axis of a rolling wheel of radius r_w where $d>r_w$. *Figure 5.10* illustrates such curves for different initial conditions and shows the first focusing point at $n=1$.

Designs of mass spectrometers with suitable fields to produce cycloidal ion paths of the prolate form have been developed by several workers[10,11,12] Aimed at compact spectrometers for residual gas analysis at low pressures, the designs are similar and a typical example is illustrated in *Figure 5.11*. The uniform electric field over the required region is provided by stacked plates accurately positioned and insulated with ceramic spacers (not shown for

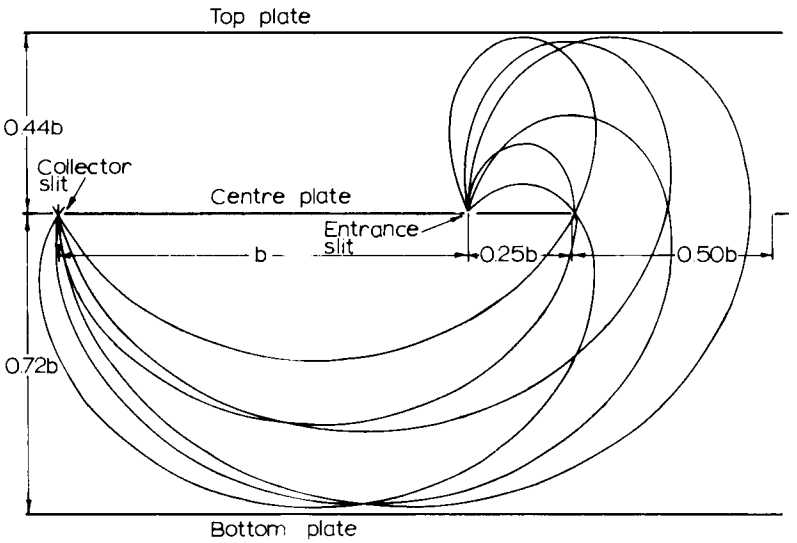


Figure 5.10 Typical prolate cycloidal trajectories used in a cycloidal mass analyser

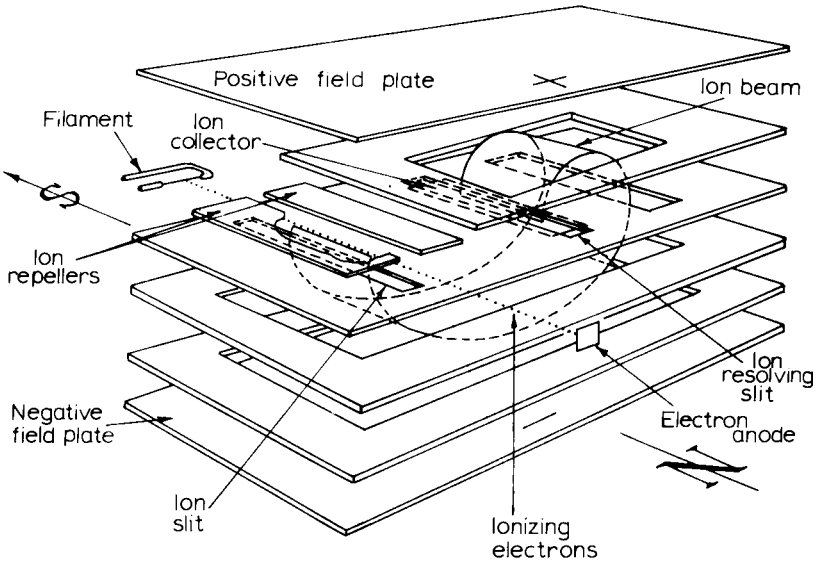


Figure 5.11 Cutaway view of a cycloidal mass spectrometer, for example by Perkins and Charpentier[10]

clarity). The plates are cut away to allow passage of the ion beam and each plate is separately connected to a supply voltage. The design necessitates a large number of components, which introduces large surface areas and trapped-gas volumes to be outgassed. It also requires a large number of electrical 'feed-throughs'.

A design to overcome these disadvantages was introduced by Andrew[13] using a fine resistance wire wrapped around four corner insulating pillars in a helix to form a potentiometer between the centre plate and two condenser plates. The design is shown in *Figure 5.12*. The main disadvantage of the design was the limitation of the voltage across the wire and therefore limitation of mass range for electrical scanning.

Because of the double focusing effect, the cycloidal mass spectrometer offered potentially a better performance in terms of resolution and sensitivity over the 180° magnetic sector type analyser. However, to obtain significant improvement, the uniformity of the electric and magnetic fields must be very good and the improvements attained in practical designs were somewhat outweighed by the more complex and expensive structure required. Although a few manufacturers offered commercial instruments based on the cycloidal mass analysing system in the 1960s, most of them have now been withdrawn from the market.

5.5 Dynamic spectrometers

5.5.1 Time-of-flight instruments

The application of the time-of-flight principle to mass separation was first suggested by Stephens[14] and was the earliest type of dynamic spectrometer.

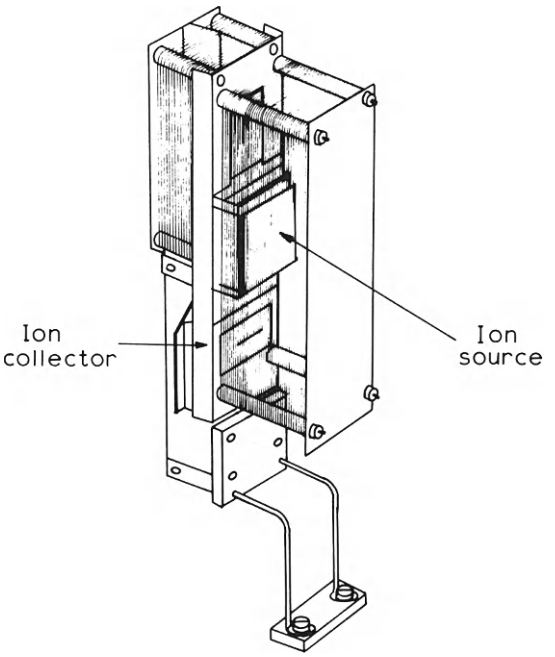


Figure 5.12 Andrew's[13] design of cycloidal spectrometer using wire helices for obtaining the correct electric field distribution

In principle, ions are accelerated through a potential V to acquire a velocity v where

$$v = \sqrt{\frac{2eV}{m}} \tag{5.11}$$

The accelerated ions then enter a field-free space and travel a distance d before arriving at a collector. The time-of-flight of the ions in the field free space is given by

$$t = \frac{d}{v} = d\left(\frac{m}{2eV}\right)^{1/2} \tag{5.12}$$

Thus, if only a small bunch of ions is allowed to leave the ion source, by pulsing the accelerator electrode, for example, then at the collector this will be divided into a number of smaller bunches separated by short time intervals. Each of these smaller bunches will correspond to a particular mass-to-charge ratio of the gas species in the original sample. The resolution depends on a mono-energetic ion source and a time-of-flight which is long compared with the time of the pulse producing the original ion bunch. The time-of-flight depends on the accelerating voltage and the length of the drift tube. The pulse time depends on the time required to build up a sufficient ion current and the time constant of the collector, but clearly it should be as short as possible if the drift tube is to be of a sensible length.

Several designs of mass spectrometer working on the principle of the time-of-flight were described in the early 1950s. However, they tended to have poor resolution, typically only just resolving masses 2 amu apart at mass 20, or they needed a long drift tube of up to 1 m in length. Later developments gave improved performance and Wiley and McLaren[15] describe an instrument capable of unit mass resolution well above 100 with useful resolution up to 300 amu using a drift tube of 40 cm. In their paper they present a detailed account of the design criteria for a time-of-flight mass spectrometer. They point out that the position of the initial ion formation as well as thermal velocities affect the resolution. The former dictates the time the ions are in the accelerating field. By using a source with two accelerating regions where distances and accelerating potentials are optimized and by introducing a time lag between production of the ions and the application of the accelerating pulse, the initial 'bunch' had a much narrower energy spread than had hitherto been attained. The ion beam was controlled by a 100 V pulse of 0.1 to 1 μs duration. To detect and amplify the ion current a multiplier with a fast rise time was required and a magnetic electron multiplier was designed for this purpose. The essential arrangement of the mass spectrometer is shown schematically in *Figure 5.13*. A further improvement in the performance of a time-of-flight analyser was claimed by Poschenrieder and Oetjen[16] by combining the linear drift space with a magnetic or electrostatic sector field. The effect of such a field is to reduce the dispersion of the analyser due to initial ion velocity in the ion source. They calculated that a spectrometer with a path radius of 20 cm could yield a resolution well beyond 600 while accepting an energy spread of 10 eV.

The main advantage of the time-of-flight spectrometer is the speed with which a spectrum can be obtained, of the order of a few microseconds. It is thus

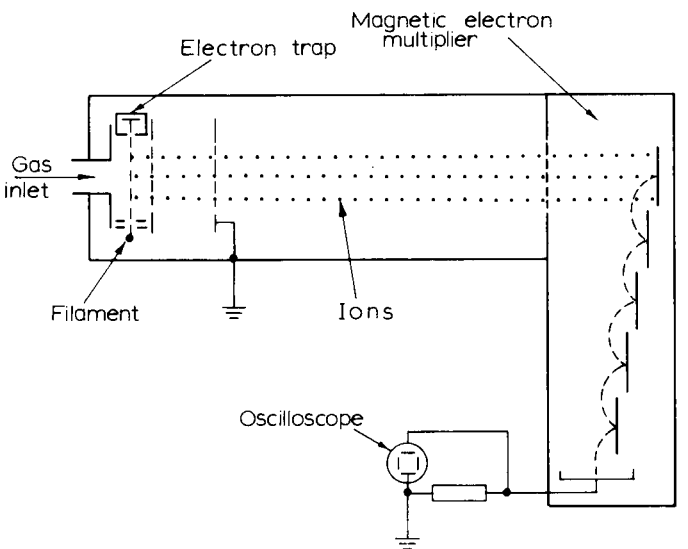


Figure 5.13 Schematic diagram of a time of flight spectrometer according to Wiley and McLaren[15]

very convenient for examining dynamic systems. Also, because of the short time, stabilization of the electron beam, etc. is not required nor is there any tight tolerance placed on mechanical alignment. The need for a fast multiplier and a wide band amplifier, however, makes it difficult to design an instrument which can be cost competitive with other types and it has therefore found use only where the fast scan time is essential.

An extension of the time-of-flight principle is to contain the ions in a magnetic field at right angles to their direction, so that instead of drifting along a tube they orbit round and round in a circular path. If the ions enter the orbit with the same momentum, then they will all orbit round the same circle, the radius being given from Equation (5.1).

$$r = \frac{mv}{eH} \tag{5.13}$$

The orbital period, however, will be given by

$$t = \frac{2\pi}{\omega_c} = \frac{2\pi m}{eH} \tag{5.14}$$

where ω_c is the cyclotron angular velocity $= v/r$.

Thus the orbital period for different masses will be different and cause the bunch to split up. By measuring the orbital period using an electronic timing circuit a mass spectrum can be obtained.

Another technique on similar lines was employed by Smith[17] whereby the ions were accelerated into the first half orbit and, as they passed through a slit system, were pulsed to give them a different orbit. A second pulse was applied later to orbit the ions to the collector. The first half orbit selected the ions as in a 180° magnetic sector analyser and the pulse technique acted as a further filter. High resolutions can be obtained by this method. These cyclotron-type or orbiting time-of-flight spectrometers, however, were not commercially exploited.

5.5.2 The omegatron

The omegatron has much in common with the cyclotron-type spectrometer described above. From Equations (5.13) and (5.14) we see that for an ion orbiting in a magnetic field the radius of the orbit is dependent on the ion velocity, but the time-of-flight or orbital period is independent of velocity. In the omegatron a r.f. field is applied such that those ions whose orbital period is resonant with the r.f. frequency gain energy and thus velocity, and spiral outwards to arrive finally at a collector. Other ions of different mass do not gain as much energy and their orbits do not expand sufficiently to allow them to reach the collector. The mass range is scanned by varying the r.f. frequency.

The principle was first applied to a mass spectrometer by Sommer *et al.*[18], and their design is illustrated in *Figure 5.14*. Ions are produced within the analyser box by an axial electron beam (in the direction of the magnetic field). A positive potential is applied to the guard rings to produce a field which prevents the loss of ions in the axial direction so that the r.f. field can act on the ions for a greater number of cycles. The r.f. voltage is applied to the two parallel end plates and also to the guard rings through a suitable voltage divider to establish a uniform r.f. field across the box (*Figure 5.14(b)*).

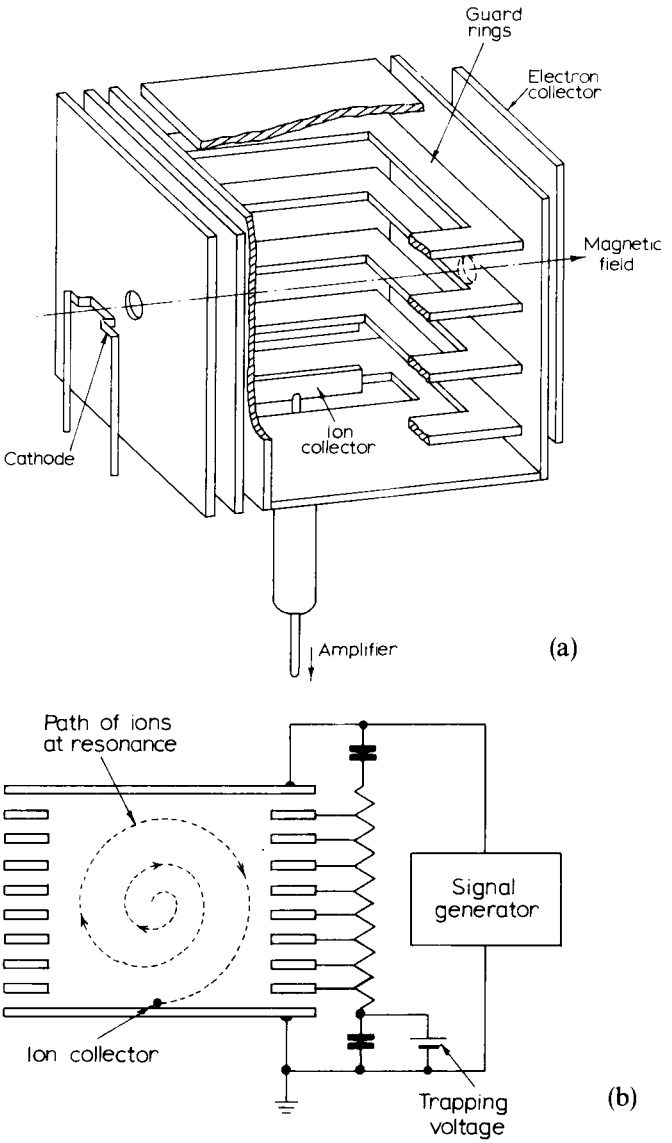


Figure 5.14 The Omegatron according to Sommer *et al*[18]. (a) Cutaway view of the structure; (b) Electrical connections etc., showing the ion path

Analysis of the motion of a charged particle, starting from rest under the influence of an r.f. field, $E = E_0 \sin \omega t$, applied at right-angles to a steady magnetic field, H, shows that the particle will describe a spiral path with a radius given by

$$r = \frac{E_0}{H(\omega - \omega_c)} \sin \frac{(\omega - \omega_c)}{2} t \tag{5.15}$$

Thus the radius of the spiral will 'beat', going through successive maxima and

minima. As ω approaches ω_c the maximum value of r increases until, at resonance, Equation (5.15) reduces to

$$r=\frac{E_0 t}{2H} \tag{5.16}$$

and there is no limitation to the maximum value of r reached by the ion. If a collector is placed at distance R from the centre of the electron beam, then only ions with $\omega-\omega_c<E_0/RH$ will reach the collector. This gives a fairly sharp resonance. If the resolution is taken as $M/\Delta M=\omega_c/2(\omega-\omega_c)$ then

$$\frac{N}{\Delta M}=\frac{\omega_c RH}{2E_0}=\frac{RH^2 e}{2E_0 m} \tag{5.17}$$

Thus the resolution varies inversely with the mass, i.e. decreasing for larger masses. A more detailed mathematical treatment of the ion trajectories, taking into account the initial velocity and initial position of the ion, has been given by Berry[19] and Brubaker and Perkins[20]. Putting values to Equation (5.17) of $H=4000$ gauss and $E_0=1$ V/cm, it is found that unit mass resolution can be obtained up to $M\sim 30$ with $R=1$ cm. Thus the analyser can be very compact and easily outgassed, making it very suitable for ultrahigh vacuum application. The early instruments were rather restricted in resolution and sensitivity. The latter mainly because the electron current had to be kept low at around 10^{-5} A. Also they tended to be critical on magnet alignment for optimum performance.

Improvements in design have been described[21,22,23] which have resulted in an acceptable performance, particularly in terms of resolution. From Equation (5.17) it is evident that the resolution increases with decreasing r.f. field. However, the minimum field is limited, first by the need to impart sufficient energy to the ions and second by the fact that at low fields the effect of contact potential becomes important. Slight changes in contact potential, due, for example, to contamination of electrode surfaces, upsets the fields and affects the stability. The improved gauges employed platinum electrodes to overcome the latter problem. The instruments were also simplified, for example the guard rings were omitted and side walls inserted which were electrically connected to one of the end plates[23].

Omegatrons were used with considerable success in the 1960s and early 1970s and were commercially exploited. A typical commercial instrument would have a mass range of 1–250 amu with unit resolution at 1% valley up to 40 amu. A mass spectrum for such a gauge is shown in *Figure 5.15.*

Although the gauge head is compact, to obtain the required magnetic field of 4000 gauss required a large magnet (30–40 kg weight) and the control box is fairly complex and expensive. Also the sensitivity is limited, since a multiplier cannot be incorporated. For these reasons it has been generally superceded by other types.

5.5.3 RF mass spectrometer

Another form of spectrometer, using a radio frequency field, has been developed for residual gas analysis. Known simply as a radio frequency spectrometer, the ions take a linear path through a succession of r.f. modulator

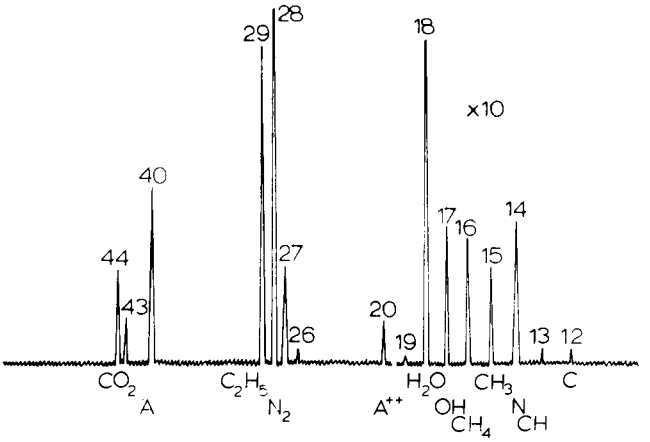


Figure 5.15 Typical mass spectrum from the Omegatron gauge

stages. Ions having a designated velocity are preferentially accelerated to overcome a barrier field at the end of the r.f. stage and are collected, whilst ions with other velocities are rejected. The first spectrometer of this type was described by Bennett[24] and is illustrated schematically in *Figure 5.16*. Electrons from the filament are drawn to the first grid to produce ions which are then accelerated through a system of three parallel grids placed a distance d apart. An r.f. field of $E_0 \sin(\omega t + \theta)$ is applied between the first two grids and a field $-E_0 \sin(\omega t + \theta)$ between grids 2 and 3. Ions arriving at the correct phase of the r.f. voltage gain energy from the field if their time-of-flight through the grids corresponds to one cycle of the voltage, i.e. if

$$\frac{2d}{v} = \frac{2d}{\sqrt{\frac{2eV}{m}}} = \frac{1}{f} \tag{5.18}$$

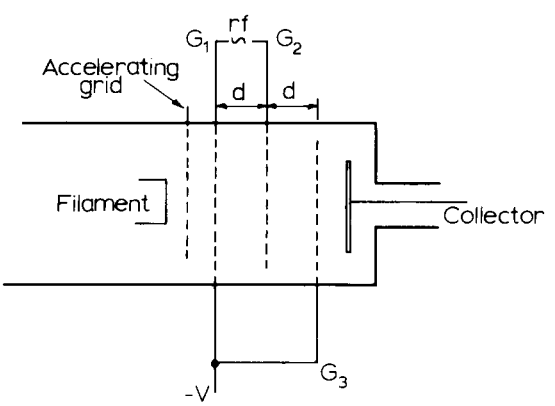


Figure 5.16 Schematic diagram of a three-grid r.f. mass spectrometer

whence

$$m = \frac{eV}{2d^2f^2} \tag{5.19}$$

A retarding potential is applied to the collector to turn back all those ions that have not acquired extra energy in the field.

Equation (5.18) represents the maximum gain of a resonant ion and does not take into account initial velocity spread or the gain in velocity in the field. A comprehensive analysis of Bennett's instrument has been given by Shcherbakova[25]. Bennett describes a practical instrument with up to three stages (of three grids each) having a drift space between each stage. Since the mass collected depends on the accelerating voltage and the r.f. frequency, either can be varied to produce a mass spectrum.

Several other workers have described similar forms of r.f. mass spectrometer for residual gas analysis. Most of the later designs are single stage but with a large number of modulator grids, with alternate grids interconnected. Varaldi *et al*[26] describe such an analyser which has been developed commercially using 12 modulating grids. The commercial model gives a 50% valley resolution up to mass 40 with a total mass range of 2–150 amu and with a lowest detectable partial pressure of 10^{-9} Pa. The spectrometer is calibrated to measure partial pressure to an accuracy of $\pm 10\%$ and current to the first accelerating grid can be used for total pressure measurement.

Robinson[27] describes a slightly different form of r.f. spectrometer based on a design by Boyd and his co-workers for plasma analysis. In this design the grids are replaced by metal cylinders 2 mm diameter and 1 mm long placed 1 mm apart. The arrangement is shown in *Figure 5.17*, together with the voltage distribution. The ions are selected by the retarding potential on double grid 2 and post accelerated to the collector by grid 3. A positive potential applied to the plate in front of the collector repels secondary electrons. Sridharen[28] has suggested that by varying the voltage of such a suppressor grid in an r.f. spectrometer the resolution can be increased by a factor of 1.3 without affecting sensitivity.

The r.f. mass spectrometer is a compact instrument requiring no magnetic field. It has a rather limited sensitivity and resolution and is best suited to industrial applications, in particular for process control applications where the partial pressure at specific mass numbers can be monitored.

A further type of r.f. mass spectrometer for residual gas analysis was devised by Tretner[29] in which ions were oscillated in a potential well set up between two parallel electrodes. The oscillation frequency is mass dependent and those ions accelerated into the system at a frequency resonant with the oscillations, build up a signal which is detected. The masses are scanned by varying the input frequency. The instrument has a rather limited resolution of around 10 amu but has found use as a monitor of residual gases.

5.5.4 The quadrupole spectrometer

The idea of the quadrupole spectrometer was first proposed by Paul and Steinwedel[30] in 1953 and was probably heralded as yet another r.f. type of spectrometer. However, its potential was soon realized and today it is

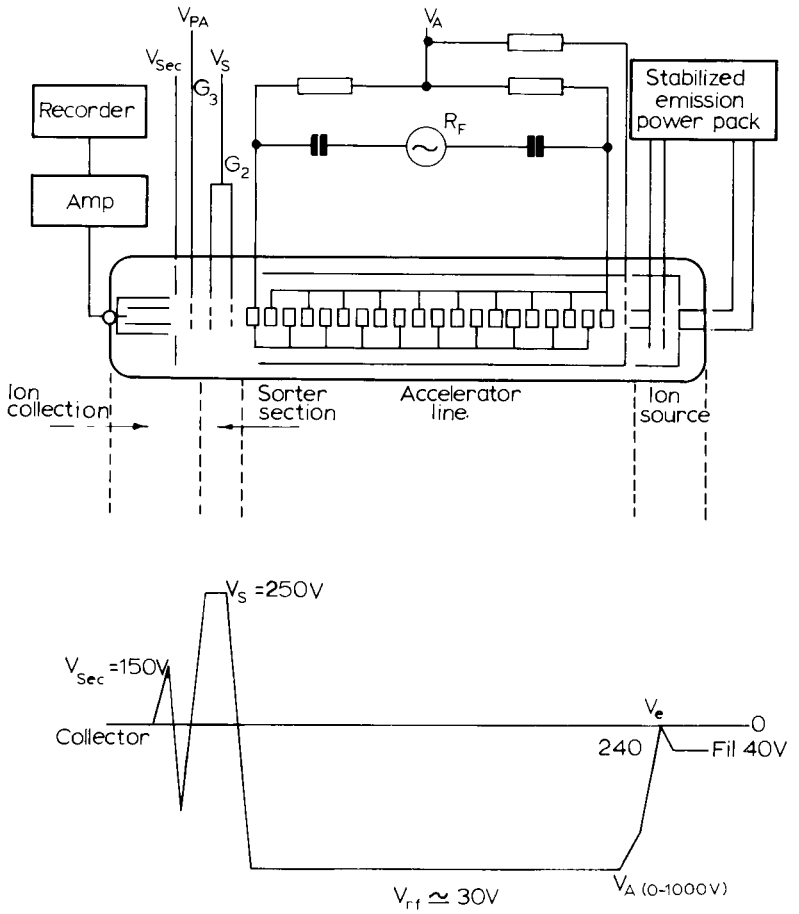


Figure 5.17 Schematic diagram of the r.f. mass spectrometer by Robinson[27] together with the voltage distribution along the ion path

established as a versatile general-purpose residual gas analyser, competing fiercely with the rival magnetic-sector type instrument.

Ideally, the quadrupole consists of four metal rod electrodes of hyperbolic cross-section, mounted as shown in *Figure 5.18*. Ions are accelerated in the z direction to pass along the space between the rod electrodes. A two-dimensional quadrupole field is set up between the rods by applying a d.c. voltage U and an r.f. component $V_0 \cos \omega t$ as illustrated, i.e. opposite rods are electrically connected and the voltages are applied across the two terminals. The potential at any point x, y within the quadrupole may be written in terms of rectilinear co-ordinates as

$$V_{xy} = \frac{(U + V_0 \cos \omega t)(x^2 - y^2)}{2r_0^2} \tag{5.20}$$

where $2r_0$ is the rod spacing.

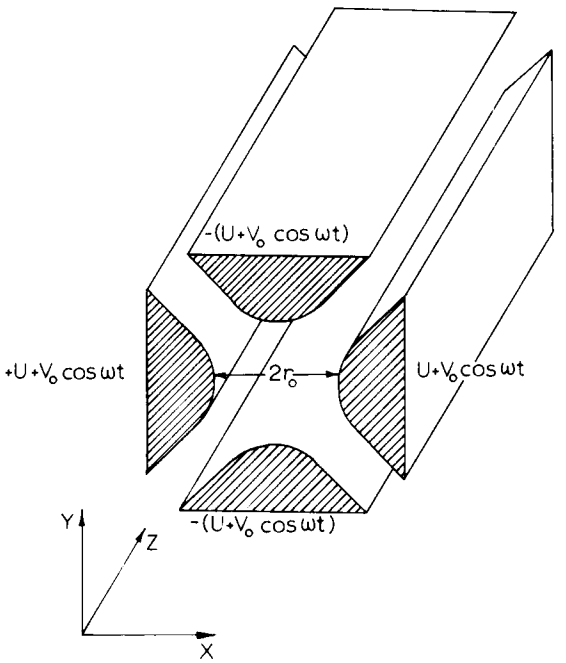


Figure 5.18 Electrode arrangement and applied potentials for a quadrupole field analyser

If two dimensionless parameters q and a are defined such that

$$q=\frac{2eV}{m\omega^2 r_0^2} \quad \text{and} \quad a=\frac{4eU}{m\omega^2 r_0^2}$$

then the equations of motion of an ion in the quadrupole field can be transformed to 'Mathieu type differential equations'.

$$\begin{aligned}&\frac{\partial^2 x}{\partial(\omega t)^2}+(a+2q\cos\omega t)x=0\\&\frac{\partial^2 y}{\partial(\omega t)^2}-(a+2q\cos\omega t)y=0\end{aligned} \tag{5.21}$$

The solutions for the ion motion in the x and y directions are independent and are oscillatory with, in general, amplitude increasing with time. When, however, q and a lie within certain discrete ranges, the amplitudes are limited in both the x and y directions and the particle performs stable oscillations within the system. The largest of these stable regions is a curvilinear triangle in the q–a plane as illustrated in *Figure 5.19*. The motion of the ions in the z direction, parallel to the electrodes, is unaffected by the fields.

By selecting the parameters a and q the ion species which remains in the stable region to pass through to the collector can be determined. In particular,

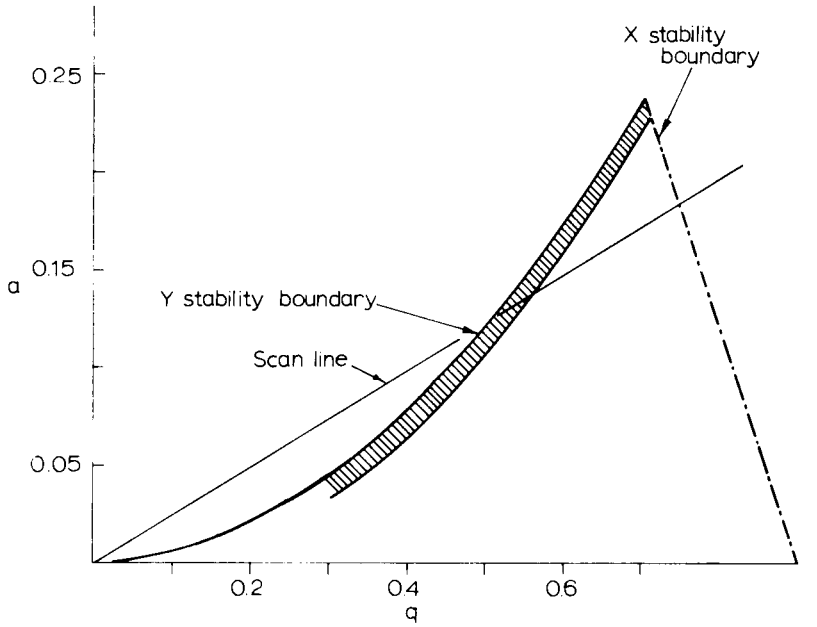


Figure 5.19 Stability diagram of a quadrupole spectrometer. (Cross hatched area represents stability area for a monopole)

if the ratio of a to q is kept constant, ions of different masses will fall on a straight line which can be made to cut the stable region near the apex as shown in *Figure 5.19*. Masses within the stable region will proceed through to the collector with typical trajectories illustrated in *Figure 5.20*, whilst masses outside will be lost by lateral deflection to the electrodes. How the line cuts the stable region determines the resolution and ion current. If it just touches the apex, only one mass will fulfil the equations and $M/\Delta M \rightarrow \infty$ but there will be no current. On the other hand, if it cuts too far below the apex, several mass numbers will lie within the stable region giving poor resolution but high current. Thus, be selecting the $a:q$ ratio, resolution can be traded for sensitivity and vice versa. Since the position of a mass number on the line is determined by the value of V_0 and U, the mass range can be scanned by holding a/q constant and varying V_0 and U simultaneously. This gives a linear scan along the line illustrated in *Figure 5.19*. The line is known as the mass scan line or operating line.

Since the oscillations are independent of the velocity in the z direction, the initial ion velocity is not critical and a simple ion source to give maximum ion current can be employed. The only criterion on ion velocity is that the ions should make several oscillation cycles before reaching the collector.

The above analysis represents an idealized model of the quadrupole mass selection system and, in practice, there are several other factors which have to be taken into account. For example, in the practical instrument mass separation depends on whether an ion is able to travel through an analyser of finite length and whether the oscillations of ions falling outside the stable

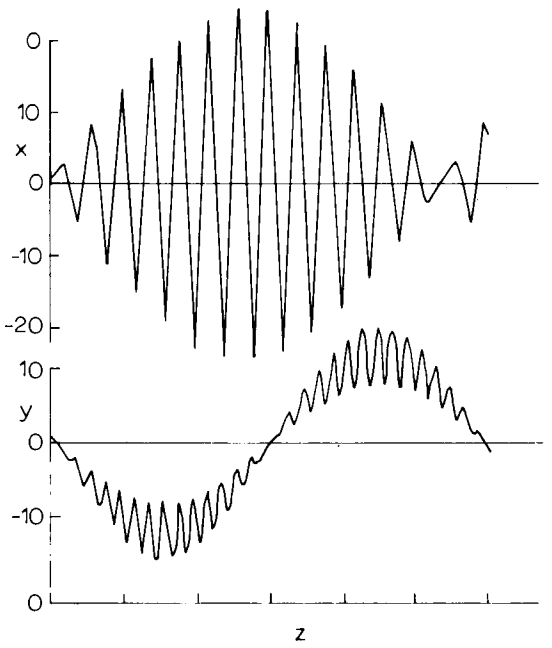


Figure 5.20 Typical ion trajectories, for a point near the tip of the stability diagram, through a quadrupole spectrometer (from Dawson[42])

region have grown sufficiently to eliminate the ions. The longer the path length, the better the resolution and the closer to the theoretical stability criterion the results are likely to be. Also, in a practical design it is usual to choose cylindrical rods[31,32], since these are easier to manufacture to the required tolerances. Provided the dimensions are suitably chosen, the field distribution for cylindrical rods approaches that of a hyperbolic quadrupole. According to Denison[33] this is attained when the radius of the rods r relative to the value of r_0 is given by $r = 1.147\ r_0$. The computed field plots for cylindrical rods are shown in *Figure 5.21*, and compared with those for hyperbolic rods.

The cylindrical rod system cannot be mathematically treated in the same way as the hyperbolic system and design of the instrument and assessment of its performance involves numerical methods employing a digital computer; normally ion trajectories are computed for a limited number of initial conditions. Computation is further complicated by the need to take into account the effect of fringing fields at the entrance and exit to the quadrupole filter. A recent theoretical approach which is not restricted on initial conditions is that known as the phase-plane approach. It was introduced by Baril and Septier[34] and has been further developed by Dawson[35]. Essentially, a plot is made of the transverse velocity $\dot{u}$ against transverse position u and an acceptance area is calculated which encloses those initial conditions which result in ion transmission under specific values of a, q and initial field phase. For a perfect field the acceptance area is a different ellipse in the x and y planes and for each initial field phase. This is illustrated in *Figure 5.22* where ellipses in the x and y direction are shown for a specific example. The acceptance area

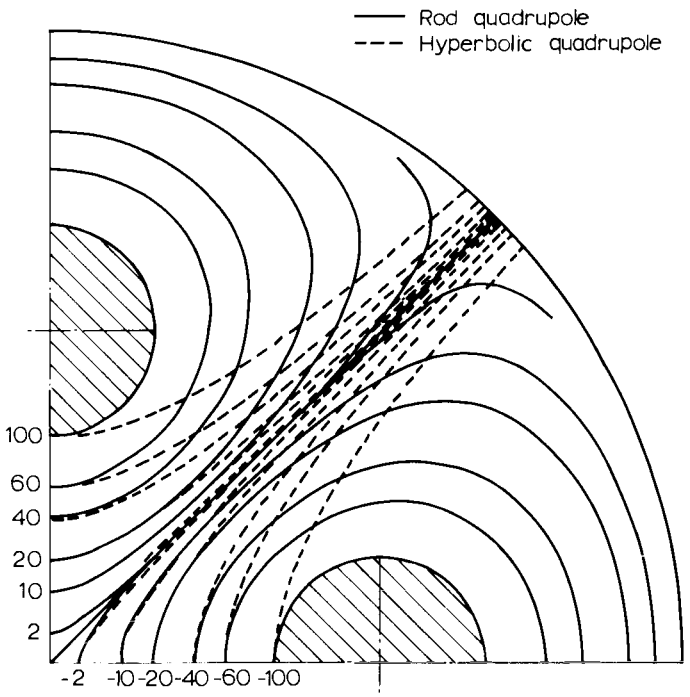


Figure 5.21 Potential distribution in a 'cylindrical rods' quadrupole compared with the 'hyperbolic rods' quadrupole

can be 'transformed' through the fringing field and indeed through any lens system to give the acceptance area for the complete instrument. By matching the acceptance with the emittance of the source the optimum performance can be obtained.

Fringing fields are a particular problem of the quadrupole since ions whose trajectories would be stable within the analyser can become unstable when they pass through the reduced field at the ends of the quadrupole rods. This can be appreciated by reference to *Figure 5.19* where the working point for a particular mass will move down the mass scan line with reducing field. Depending on the velocity, and thus the time the ion spends in the fringing field, these instabilities can cause partial rejection of the ions of the desired mass or cause them to reach the interior of the analyser under unfavourable conditions of off-axis distance and velocity direction. Thus, the sensitivity and resolution can be affected. One solution suggested by Brubaker[36] was to use an additional set of rod electrodes at the entrance to the quadrupole with potentials differeing from those applied to the main rods to keep the working point within the stable position as the ions transverse the fringing fields. To achieve this the 'entrance' quadrupole had a.c. signals applied with small or zero d.c. potential, so that the ions come first under the influence of the a.c. field and further down the analyser the d.c. field. This system is called 'delayed d.c. ramp' and can be fabricated by extending the metal rod electrodes with insulating portions. An alternative arrangement, using the same principle, has been described by Brubaker[37] in which metal 'entrance' electrodes are used

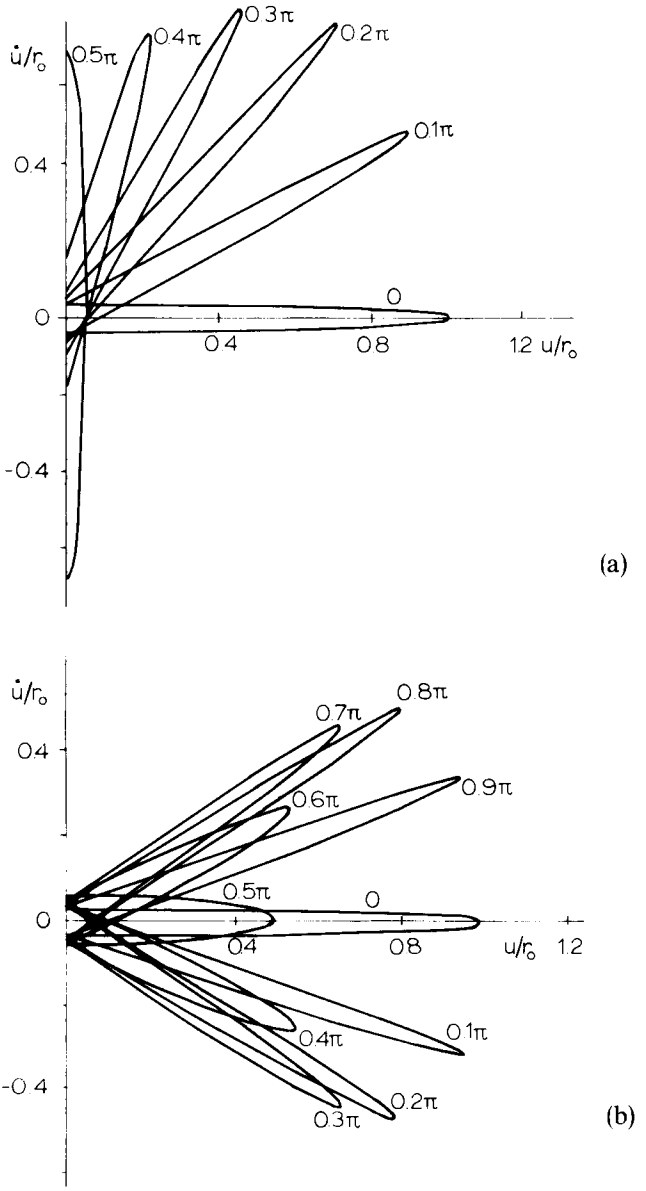


Figure 5.22 Acceptance ellipses for the quadrupole mass spectrometer $a = \pm 0.2334$ and $q = 0.706$. (a) In the x direction, initial phases from 0 to 0.5π being shown; (b) In the y direction (from Dawson[42])

with d.c. potential only applied but with a polarity opposing that of the analyser rods. In this way the d.c. field at the fringe is neutralized without affecting the a.c. fringe field to give the same delayed d.c. ramp condition. Fite[38] suggests that a similar effect can be obtained by placing a 'leaky dielectric' tube, through which the ions pass, at the entrance to the quadrupole.

The dielectric material is chosen such that it appears to the r.f. field as a dielectric but to the d.c. field as a conductor. Fite suggests a number of materials which could fulfil the conditions but even with the most suitable material, his experimental results were rather inconclusive.

A different approach has been proposed by Brinkmann[39] to overcome the problem of fringing fields. He applied an a.c. field only to the rod electrodes using the analyser as a high pass filter, i.e. $a = 0$ in Equation (5.21). In this mode a range of ion masses passed without transmission loss through the rod assembly. However, at the exit the ion energies are strongly influenced by the axial component of the exit fringing field. In particular, ions well off-centre gained more energy than those passing through the centre. Thus, those ions near the edge of the stable region gained more energy than those well inside it. Mass separation could be obtained by applying a retarding field to the collector. In practice the retarding field was held constant and the mass range scanned by varying the r.f. voltage amplitude. Brinkmann obtained an increase of sensitivity of ten times that attained in the normal mode.

Holme[40] has shown that the method can also be used to give an increased resolution. A slight disadvantage of the system is that extra 'satellite' peaks are observed. A more extensive study by Holme *et al.*[41] demonstrated the advantages of the system for a small quadrupole designed at Liverpoool University. In particular they found there was less dependence on mechanical alignment. *Figure 5.23* is taken from their paper and shows the enhanced resolution for Xenon and also the satellite peaks. For a more detailed discussion on the quadrupole spectrometer theory and characteristics, the reader is referred to the excellent book on the subject edited by Dawson[42].

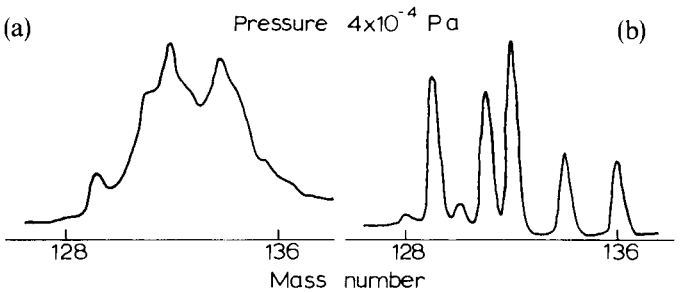


Figure 5.23 (a) Mass spectrum from a small quadrupole working in conventional mode; (b) Mass spectrum from the same quadrupole using the retarding field mode (i.e., a.c. potential only on the electrodes). According to Holme *et al.*[41]

The early quadrupole spectrometers were relatively large, with rods 20 cm or so in length and with an overall diameter of up to 10 cm. They were also expensive. Because of fringing field and alignment problems thier mass range was not particularly good, typically 1–100 amu, and most were fitted with multipliers to enhance their sensitivity. They frequently suffered from problems of drift, due to distortion of the rods during baking cycles or contamination upsetting the contact potentials. As a result, recalibration was required at frequent intervals.

Better construction and choice of materials for the rods and supports over the last two decades have resulted in vast improvements in the performance. Today, instruments of a similar size to those early spectrometers offer a mass range of 1–1000 amu with a possible sensitivity of 10^{-5} A Pa^{-1} and with scanning speeds up to 10 ms. More important, smaller instruments have been constructed which are much more economic but which still have adequate mass range and sensitivity for most residual gas analysis applications.

The first of these small instruments was described by Bargery and Ball[43] in 1968. It had rods 15 cm in length and gave unit mass resolution up to 100 amu with a sensitivity around 10^{-6} A Pa^{-1}. However, from the study of small quadrupoles at Liverpool University, Holme *et al.*[44] showed that the rod length could be reduced to 5 cm with very little reduction in performance. They investigated especially the effect of errors in analyser construction and in the applied potentials by deliberately introducing errors, such as off-setting the source assembly and superimposing noise on the supply potentials. They found that provided the construction and assembly of the analyser are to good machine tool standards and the potentials are well stabilized, the only factor limiting the maximum resolution under any set of operational conditions was the time spent by the ions in the quadrupole field, or more precisely, the number of oscillations they make. This is a function of V_{rf}/V_i where V_{rf} is the r.f. voltage and V_i the ion acceleration voltage. V_{rf} should, therefore, be as high as possible and V_i as low as possible. V_{rf} was limited by breakdown considerations to a maximum of 1 kV and V_i was optimized at about 10 V. The frequency was set at 4 MHz to cover 0–50 amu range and 2 MHz for 0–200 amu range, the latter at a reduced sensitivity.

Commercial instruments of similar dimensions are now on the market. Typically they cover a 1–50 amu with 10% valley resolution with a minimum detectable partial pressure of around 10^{-9} Pa. They represent compact bakable partial pressure gauges offering extra information on the vacuum environment over an ion gauge at very little extra cost. *Figure 5.24* illustrates one such instrument marketed by Hidden Analytical Ltd. This has a mass range of 2–100 amu with a minimum detectable partial pressure of 10^{-9} Pa.

An interesting design in this small analyser category is that described by Reich[45] in which the quadrupole analyser consists of a single piece ceramic body upon which the electrodes are coated, shaped to give the hyperbolic quadrupole cross-section distribution down its centre. The advantage of this construction is its stability with temperature, allowing measurements to be made during bakeout at 200°C.

A modification of the quadrupole which has been progressed by one manufacturer is the monopole. First described by von Zahn[46] in 1963 it is essentially a quarter of a quadrupole consisting of one cylindrical rod and an angle plate, as illustrated in *Figure 5.25*. The potential of the plate is earth and the rod is held at $-(U + V_0 \cos \omega t)$. The field, in this case, at point xy between the electrodes is

$$X_{xy} = \frac{(U + V_0 \cos \omega t)(x^2 y^2)}{r_0} \tag{5.22}$$

i.e. twice the field of the quadrupole (Equation (5.20)).

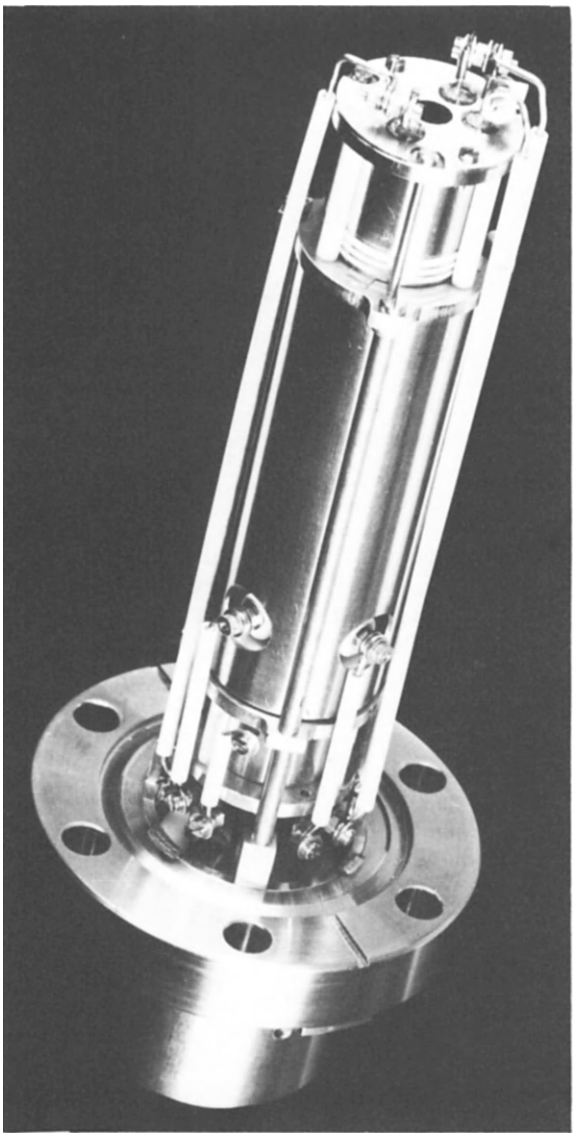


Figure 5.24 Photograph of the small quadrupole Hal 100 mounted on a 70 mm diameter ConFlat® flange. (Courtesy Hidden Analytical Ltd)

By defining

$$a_m = \frac{8eU}{m\omega^2 r_0^2}, \qquad q_m = \frac{4eV_0}{m\omega^2 r_0^2}$$

the motion of the ions in the monopole can be expressed by the same Mathieu equations as for the quadrupole (Equation (5.21)). However, the monopole differs from the quadrupole in its mode of operation. Clearly, if the ions are not

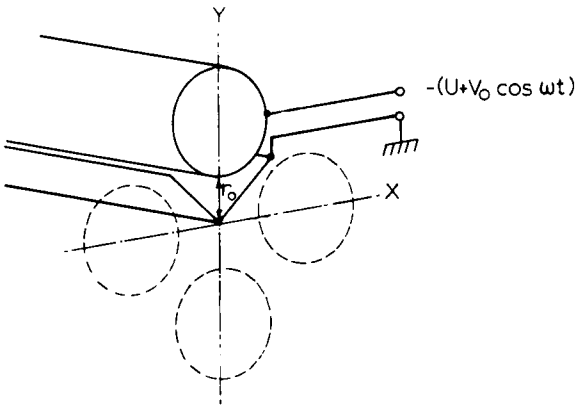


Figure 5.25 Electrode arrangement and applied potentials for the monopole spectrometer

to impinge on the plate, the x deflection must be less than the y deflection which, in turn, must always be positive. From the trajectories shown in *Figure 5.20* it is seen that to fulfil the latter condition the number of oscillations in the y direction must be fewer than the number in a 'beat'. Also only those ions entering in the phase which gives a positive oscillation will pass through, i.e. 50% of the initial beam. These restrictions were shown by von Zahn to limit the stability area on the $q:a$ plot to a narrow band on the left hand of the stability boundary, shown as a cross-hatched area in *Figure 5.19*. As a result, the working point can be well below the apex without losing resolution. Oscillations in the x direction are not a problem since the voltages are lower due to the factor 2 in Equation (5.22) relative to (5.20) and the high ratio of $q:a$ that can now be used.

Thus, the monopole has the advantages of a simpler structure, lower working voltages and a resolution which is less critical on the U/V_0 ratio. Its main disadvantage is that the ion velocities are more critical and, therefore, a mono-energetic ion source is required. The resolution of the monopole depends on the accelerating voltage; the lower the voltage the higher the resolution. However, the converse is true of sensitivity and a compromise is required. Because of the variation of sensitivity with accelerating voltage, the instrument is not so satisfactory as a partial pressure gauge. The monopole has been less thoroughly examined than the quadrupole and consequently the effect of fringing fields, misalignment, etc. is less well known. Nevertheless, the commercial monopole produced by Veeco Instruments Inc. is a very competitive instrument offering a mass range of 1–200 amu with a unit resolution up to mass 50 with a 10% valley. The analyser tube is just over 20 cm long and the minimum detectable partial pressure is about 10^{-8} Pa.

5.6 Ion detection

Having separated the ion species, it is necessary to determine the quantity of each ion species by a suitable collecting and measuring technique.

The simplest way to detect the ion current is to mount a plate electrode biased at essentially earth potential at the exit of the analyser and to measure the current to it with a sensitive current-measuring instrument. To obtain optimum sensitivity and ensure that the indicated current is the true ion current, certain precautions have to be taken. False readings will be given by secondary electrons emitted from the plate as a result of bombardment by the ions and this must be allowed for or, better still, suppressed. This can be achieved by constructing the collector in the form of a cup to prevent electrons escaping (Faraday cup) and/or by mounting an aperture plate electrode immediately in front of the collector, negatively biased with respect to the collector to suppress the electrons by turning them back to the collector plate. The system is shown diagrammatically in *Figure 5.26*. A similar turning back of the electrons can be obtained with a magnetic field and for the 180° magnetic-sector spectrometer the suppressor plate may not be necessary. There is also a possibility that ions with a higher mass than the selected value will lose energy by collision and also be collected.

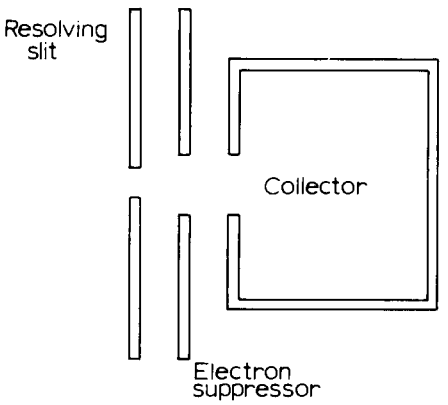


Figure 5.26 Schematic diagram of a Faraday cup ion collector arrangement

To minimize this effect, an ion suppressor plate is often inserted between the slit and the electron suppressor plate. Operated at a potential near that of the ionization region, it provides a retarding field through which scattered ions cannot pass. It has the effect of 'sharpening' the resolution, and control of its potential can be used for this purpose.

The sensitivity depends on the minimum current that can be measured and over the years that mass spectrometers have been used, current measuring instruments have been developed continuously, from quadrant electrometers, through valve amplifiers with high impedance electrometer tubes and vibrating reeds at the input, to the present solid state amplifiers employing FETs or MOSFETs at the input stage.

The minimum current which can be measured with such an amplifier is around 10^{-15} A, provided the noise level is kept to a minimum. To reduce spurious signal pickup, a main cause of noise, it is important to employ short leads and good screening. Ideally the input stage of the amplifier should be

mounted on the gauge head with the leads to the collector of no more than a few centimetres. The amplifier should be free from vibrations and not be subjected to temperature fluctuations. The power supplies need to be highly stabilized. For low currents, 10^{-15} A, the response time of the amplifier is typically a few seconds which means the scan rate is limited to less than 1 amu per second. Faster scans can be obtained only at the expense of sensitivity. To a first approximation the scan rate is inversely proportional to the minimum detectable signal at the amplifier. At the sensitivity of 10^{-6} A Pa^{-1}, typical of small residual gas analysers, a current of 10^{-15} A allows partial pressures of 10^{-9} Pa to be measured.

An improvement in the minimum current which can be measured and in the response time, can be obtained by replacing the plate collector by an electron multiplier. Basically the ions incident on the front plate of the multiplier produce secondary electrons which are directed down the multiplier via a series of electrodes or dynodes on which they impinge. The dynodes are coated with a high secondary emission material (secondary emission coefficient $\delta > 1$) so that the impinging electrons liberate a number of secondary electrons at each impact with the dynodes surface, building up an avalanche. The electrons are directed from dynode to dynode by the potential difference between them, and the geometry and electrostatic field ensures that the primaries and secondaries at each dynode are focused on to the next. The multiplier may have up to 15 dynodes, the number being limited by the current capacity of the last dynode, giving gains of up to 10^5–10^6. Because the 'build-up' depends on electron movement within the vacuum alone, the response time can be as low as a few microseconds.

Three types of multiplier are commonly available, namely, the discrete dynode electrostatic multiplier, the magnetic electron multiplier, and the channel electron multiplier.

The discrete dynode multiplier can take a number of forms, the three main designs are illustrated in *Figure 5.27*. In general, each dynode is maintained at a potential of $+200 \rightarrow 300$ V above the preceding dynode. This implies that a resistive divider chain must be connected to each of the individual dynodes and that, since the output is essentially earthed, the input has to be held around -3 kV, with the 3 kV across the resistor chain. The principal secondary emission material, which has been used for such multipliers, is beryllium–copper which forms a beryllium oxide surface with δ of the order of 3. Silver magnesium alloy is also used with a slightly higher δ.

In a magnetic electron multiplier, the dynodes are co-planar and the electrons take cycloidal paths from dynode to dynode in a crossed electric and magnetic field. The focusing effect of the cross field and the higher electric field used gives the electrons a faster transit time and makes this type of multiplier particularly suitable for time-of-flight spectrometers. A development of the magnetic electron multiplier has been described by Goodrich and Wiley[47] whereby the discrete dynodes were replaced by a continuous dynode. The continuous dynode consisted of a glass plate coated with a resistive layer so that a potential gradient could be applied across it to give the necessary electric field for the electron trajectories. In this way the resistor divider chain was eliminated.

The channel electron multiplier is a further exploitation of the continuous resistive dynode. The device consists of a glass tube typically about 40

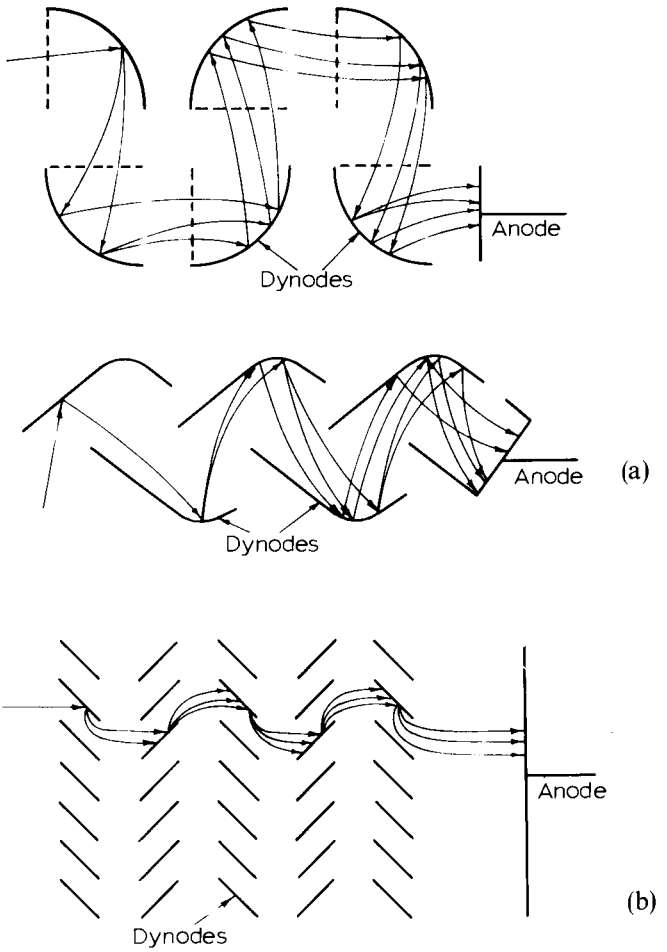


Figure 5.27 Schematic diagram of three types of discrete dynode electrostatic electron multipliers. (a) Box and grid design; (b) Focused curved dynodes; (c) Venetian blind design

diameters in length, having a high resistance. A potential of 1 to 2 kV is applied across the ends of the tube to give secondary electron trajectories as illustrated in *Figure 5.28*. Not only must the glass be resistive but it must also have a high secondary emission coefficient for its internal surface. The success of the channel multiplier has been the development of suitable glasses and their processing to give the required characteristics. The channel multipliers are usually curved to prevent the ions formed by impact ionization in the residual gas passing back down the channel and causing further secondaries at the entrance (ionic feedback). They also have a funnel-shaped entrance to collect the maximum number of ions. A typical channel multiplier for mass spectrometer applications is shown in *Figure 5.29*, together with samples of two other types of multiplier. Apart from eliminating the resistor divider chain, the obvious advantage of the channel multiplier is its small dimensions. A

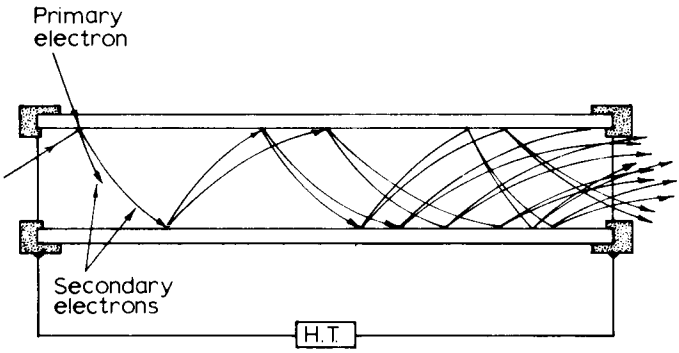


Figure 5.28 Electron trajectories in a channel multiplier

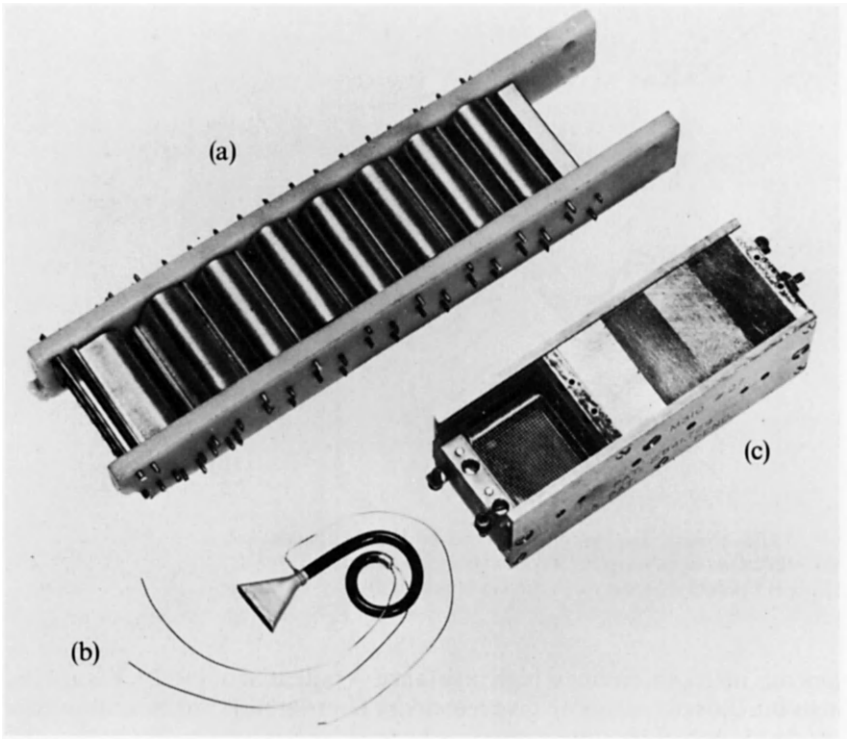


Figure 5.29 Photograph of three types of electron multiplier. (a) Discrete dynode multiplier (focused dynode type); (b) Channel multiplier; (c) Magnetic multiplier

special issue of *Acta Electronica*[48] has been devoted to the channel multiplier to which the reader is referred for further details.

It might be considered that, with a multiplication of 10^6, partial pressures well below 10^{-12} Pa could be measured with an electron multiplier. In practice, however, there is a limit due to ion statistical noise. At 10^{-12} Pa and a sensitivity of 10^{-6} A Pa^{-1} the input current would be 10^{-18} A or about 6 ions per second. Even with a slow monitoring system, the statistical arrival of

these ions will result in a noise problem. The channel multiplier can be used in an ion count mode, but the counting circuit is relatively complex and not suitable for a residual gas analyser. There is also a limitation to the total current that can be drawn from the multiplier due to space charge and the current carrying capacity of the final dynode. In general, the output current must be less than half of the current through the resistor chain or resistive layer. With these limitations, the improvement that can be obtained over a plate collector is of the order of 10^3. Nevertheless, this is a significant improvement, allowing partial pressures down to 10^{-12} Pa to be measured. Alternatively, this improved sensitivity can be traded for improved resolution. Chisholm *et al.*[49] at the author's laboratories investigated the use of a channel electron multiplier with a small 90° magnetic-sector type residual gas analyser. *Table 5.1* illustrates the gain in sensitivity and resolution that can be obtained with various potentials on the multiplier and on a grid placed in front of it, when used to measure the input ion current.

TABLE 5.1. Sensitivity and resolution of the 90° spectrometer with channel multiplier

Grid potential	*Multiplier potentials*		*Sensitivity* $A\ torr^{-1}$ *for mass 28*	*20% valley resolution (max. mass number)*
	Input	*Output*		
0	0	0	5.5×10^{-6}†	52 at 10% valley
0	0	1300	1.1×10^{-2}	56
0	0	1400	3×10^{-2}	56
0	−500	750	1.1×10^{-1}	79
0	−1200	0	2×10^{-1}	88 (77 at 10% valley)
−1200	−1200	0	1.0	62

Measured at $\sim 5 \times 10^{-6}$ Pa and 100 μA emission current

† The sensitivity was measured at the multiplier input and was more than one order down on that obtained in the original instrument with a Faraday cage

There is, however, one big snag with the use of any multiplier. The multiplication depends critically on δ and this in turn is a surface effect. Contamination of the surface by residual gases, particularly when the system is let up to air, and cleaning of the surface by baking and electron bombardment strongly affect the multiplication factor. It therefore requires constant re-calibration and will sometimes drift whilst measurements are being made. Also the secondary emission at the entrance depends on the ion mass, giving further mass discrimination. For mass spectrometers where there is a direct line of sight path between ion source and collector, the multiplier must be mounted off centre to prevent secondary electrons being produced by photons from the ion source.

Summing up, the plate detector is simple and, for many applications, adequate. The readings are reproducible but the lowest pressure that can be measured is around 10^{-10} Pa with a response time of approximately 1 amu per second. This means that the spectra must be recorded on a chart recorder. With a multiplier a lower pressure can be measured and the response time can be considerably reduced. At 10^{-10} Pa pressure, for example, a response time of 1 amu per millisecond is attainable. The output can therefore be coupled to an

oscilloscope. The output, however, is not very reproducible and may drift during operation. The multiplier system is also at least twice the price of the plate collecting system.

5.7 Calibration and performance

The usefulness of spectra obtained from a residual gas analyser depends on the calibration. As mentioned in the introduction, ionization of gas molecules can produce singly and doubly charged ions of the molecules as well as fragment ions. For example, carbon dioxide, when bombarded by electrons, will produce ions of CO_2^+, CO^+, CO_2^{++}, O^+ and C^+ in significant quantities as well as small quantities of isotope ions, giving a characteristic spectrum with mass peaks at masses 44, 28, 22, 16 and 12 respectively. Fortunately, however, the relative abundance of the various ions for any one gas is more or less constant for any given set of conditions and is known as the cracking pattern of the gas. In general, these cracking patterns are independent of pressure but they do depend on the electron energy and the gas temperature and also on the mass discrimination. The temperature affects the dissociation probabbility and thus more fragment ions are produced at higher temperatures. The electron energy determines the ratio of singly to multi-ionized molecules. The mass discrimination. The temperature affects the dissociation probability of scanning, etc., will obviously affect the peak heights for the various masses. Because of this latter effect, different types of mass spectrometer can exhibit significant differences in cracking pattern for the same gas, but instruments of the same type tend to show only minor variations.

Cracking patterns for most gases have been investigated for the magnetic sector type instruments and can be found tabulated in any mass spectrometer data book. Similar information is being built up for the quadrupole. Conventionally the largest peak is scaled as 100, and is used as the base peak for sensitivity measurements, etc. *Table 5.2* gives some typical cracking pattern data in a magnetic sector instrument for the gases likely to be met in vacuum systems. Knowing the cracking pattern allows the spectrum to be analysed. For example, if there is a mixture of CO_2 and N_2 present in the system, both will contribute to the peak at mass 28. However, by measuring the peak at mass 44, which is unique to the CO_2, the contribution of CO^+ to the mass 28 peak can be deduced and thus the relative height of the N_2^+ peak at 28 determined.

Whilst the cracking pattern data identifies the residual gases present and enables their relative abundance to be estimated in terms of the peak heights, the actual partial pressure requires knowledge of the sensitivity of the instrument to the various gases. Thus calibration of the residual gas analyser implies measuring cracking patterns and sensitivities, ideally for all the gases likely to be present. In practice this would prove an impossible task and normally the instrument would be calibrated for one or two gases and the performance for other gases deduced from the results and published data.

The calibration apparatus needed is similar to that outlined in Section 4.4 for total pressure gauges. The instrument can be calibrated in either a sealed-off system or by the orifice method. The pressure is measured using the base peak for each gas against a standard total pressure gauge calibrated for that

TABLE 5.2. Mass spectra cracking patterns

	Methane	*Ammonia*	*Water*	*Neon*	*Acetylene*	*Ethylene*	*Carbon monoxide*	*Nitrogen*	*Ethane*	*Nitric oxide*	*Methyl alcohol (methanol)*	*Oxygen*	*Hydrogen sulphide*	*Argon*	*Propene*	*Propane*	*Carbon dioxide*	*Nitrous oxide*	*Acetaldehyde*	*Ethyl alcohol (ethanol)*	*Nitrogen dioxide*	*Formic acid*	*N. butane*	*Acetone*
Molecular weight	16	17	18	20	26	28	28	28	30	30	32	32	34	40	42	44	44	44	44	46	46	46	58	58
Formula	CH_4	NH_3	H_2O	Ne	C_2H_2	C_2H_4	CO	N_2	C_2H_6	NO	CH_4O	O_2	H_2S	A	C_3H_6	C_3H_8	CO_2	N_2O	C_2H_4O	C_2H_6O	NO_2	CH_2O_2	C_4H_{10}	C_3H_6O
Mass																								
2	3.0		0.7										0.2											
12	2.4				2.5	2.1	4.5									0.4	6.0					3.3		
13	7.7				5.6	3.5										0.5	0.1					2.9		
14	18.6	2.2			0.2	6.3	0.6	7.2	3.4	7.5					3.9	2.5		12.9			9.6			
15	35.8	7.5							4.6	2.4					5.9	3.9		0.1					5.3	
16	100.0	80.0	1.1				0.9			1.5		11.4					8.5	5.0			22.3	5.2		
17	1.2	100.0	23.0																			17.1		
18		0.4	100.0								1.9									5.5				
19			0.1																	2.3				
20			0.3	100.0										10.7										
21				0.3																				
22				9.9													1.2							
24					5.6	3.7													1.6					
25					20.1	11.7			4.2							0.7			4.8					
26					100.0	62.3			23.0						11.3	7.6			9.1	8.3			6.3	8.3
27					2.8	64.8			33.3						38.4	37.9			4.5	23.9			37.1	8.0
28					0.2	100.0	100.0	100.0	100.0		6.4					59.1	11.4	10.8	2.7	8.9		17.2	32.6	
29						2.2	1.1	0.8	21.7		64.7					100.0	0.1	0.1	100.0	23.4		100.0	44.2	4.3
30							0.2		26.2	100.0	0.8					2.1		31.1		6.0	100.0	1.6		
31										0.4	100.0							0.1		100.0				
32										0.2	66.7	100.0	44.4											
33											1.0	0.1	42.0											
34												0.4	100.0											
35													2.5											
36													4.2	0.3		0.4								
37															13.4	3.1								2.1
38														0.1	20.3	4.9								2.3
39															74.0	16.2							12.5	3.0
40														100.0	29.9	2.8								
41															100.0	12.4			3.9				27.8	2.1
42															69.6	5.1			9.2	2.9			12.2	7.0
43																22.3			26.7	7.6			100.0	100.0
44															26.2	100.0	100.0	45.7				10.0		
45															0.8	1.3	0.7		34.4			47.6		
46																0.4	0.2		16.5		37.0	60.9		
47																								

gas and the relative height of the subsidiary peaks measured to obtain the cracking pattern.

To ensure that the sensitivity and cracking patterns can be extrapolated over the ultrahigh vacuum pressure range, calibrations down to ultrahigh pressures are desirable. The equipment, therefore, should be compatible with ultrahigh vacuum technology with a low background pressure well into the ultrahigh vacuum region. On the other hand, experience with most spectrometers has shown that the ion current is a linear function of pressure to within 1% for pressure below 10^{-4} Pa, i.e. the sensitivity is constant with pressure and that residual currents are not normally a problem.

In general the sensitivity for a particular gas, for example nitrogen, can be determined for an individual instrument to about $\pm 4\%$. However, it is essential to use the instrument under exactly the same conditions as laid down in the calibration. Changing the filament, or re-aligning the magnets, etc. can upset the calibration and contamination of electrodes can also cause deviations. Some instruments are sold as partial pressure gauges and are pre-calibrated. For a particular type of spectrometer the calibration may be within $\pm 10\%$, but this depends on the type of gauge and, in general, a check calibration is required. For this purpose it is useful to use an inert gas and compare the peak of the singly and doubly ionized atoms. For example, the argon peak at mass 40 is about 10 times the peak at 20 for a magnetic sector instrument and about six times for the small quadrupole.

The relative sensitivity for other gases can be determined by calibration to a similar accuracy or, if the manufacturers' figures are taken, then $\pm 10\%$ accuracy can be assumed. For spectrometers fitted with multipliers the same degree of accuracy can only be obtained by frequent calibration.

In some instruments the total pressure can be measured by extracting a small known fraction of ion current before mass separation. As with the ion gauge, the pressure measured in this way will be the nitrogen equivalent total pressure. The true pressure can be obtained by summing the partial pressures of the constituent gases and on this premise the mass spectrometer offers a more reliable total pressure measuring device than the ion gauge. However, this will only be true if the interpretation of the spectra is correct, and the cracking patterns and sensitivities are accurately known.

Interpretation of the spectra is a particular problem when the cracking patterns overlap. An example of this difficulty was demonstrated by Fite and Irving[50] using a high resolution instrument. In their instrument the mass 28 peak could be resolved into a triplet of CO^+ (at 27.99492 amu), N_2^+ (at 28.00614 amu) and $C_2H_4^+$ (at 28.03130 amu). They showed that interpretation of the conditions in a vacuum system based on a small residual gas analyser had been incorrectly deduced and that, under certain circumstances, the CO^+ peak could be greater than the N_2 peak at mass 28, in spite of the expected higher abundance of N_2.

Another factor which may affect the interpretation of partial pressures from the cracking pattern is the interaction of active gases with the heated filament of the ion source. Breth *et al.*[51] demonstrated this by looking at the cracking pattern of CO_2 under different pump speed conditions in a quadrupole spectrometer. Their results are illustrated in *Figure 5.30*. The lower value of the effective pump speed makes the reaction of the admitted CO_2 with the hot filament more probable, increasing the CO content in the spectrometer. Breth

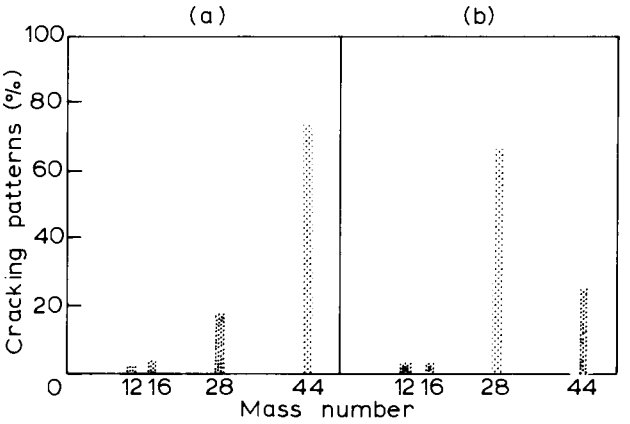


Figure 5.30 Cracking patterns for CO_2 measured with two values of effective pump speed s_{eff}. (a) $s_{eff} = 10^{-2}\,m^3s^{-1}$; (b) $s_{eff} = 8 \times 10^{-4}\,m^3s^{-1}$, according to Breth *et al.*[51]

et al. also showed that desorption effects and possible charging up of the analyser rods could introduce inaccuracies of up to 20% in partial pressure measurements for gas mixtures having common mass to charge ratios.

Summing up, the residual gas analyser has moved from the expensive luxury instrument requiring expert handling and interpretation, to an economic and reliable gauge, which is relatively easy to use and which will give considerably more information about the vacuum environment than can possibly be ascertained from total pressure gauges.

Several types of residual gas analysers have been developed, each having advantages and disadvantages. However, over the last decade several types have fallen by the wayside and today only two dominate the market, the magnetic sector spectrometer and the quadrupole.

The least expensive and simplest type is the 180° magnetic sector instrument with a beam radius of about 1 cm. It is reliable and reproducible and has been developed with suitable control circuits to cover most of the application requirements for the ultrahigh vacuum engineer. Its competitor, on price, the 5 cm rod quadrupole, has the advantage of no magnetic field and a nude ion source, which can be inserted into the vacuum system. It is also easier to outgas, having no magnets to remove. However, its resolution is poorer at lower mass numbers, 1–20 amu, and it is not as good on stability and reproducibility. Since the resolution of the quadrupole is proportional to the mass, whilst that of the magnetic-sector instrument is independent of mass, the resolution at higher mass numbers is better in the quadrupole. On the other hand, the resolution at low amu is important in ultrahigh vacuum applications, and in general, a larger quadrupole is required for similar performance, which will probably be more expensive.

Similar arguments apply to the larger higher resolution instruments, where the 60° or 90° magnetic-sector instruments compare with quadrupoles having rod lengths of the order of 20 cm. The quadrupole is more compact but the magnetic sector type still seems to be preferred where quantitative gas analysis data are required.

It is clear that improvements are still being made to the quadrupole gas analyser and it could be that, in the future, they will oust the magnetic sector instruments from the market. In the meantime, choice rests on the applicational requirements and sometimes on personal preference.

The electron multiplier has also been improved over the years, offering a useful addition to the gas analyser, extending the lower pressure limit by two or three orders and/or offering a faster scanning rate. They are now compact, readily outgassed components which, although not ideal, give acceptable stability and reproducibility for most applications. On many of the higher resolution instruments they are fitted as standard.

5.8 References

1. NIER, A. O., *Rev. Sci. Instrum.*, **11**, 212, (1940)
2. BARNARD, G. P., *Modern Mass Spectrometry*, Inst. Phys. London, (1953)
3. PITTAWAY, L. G., *Philips Res. Rep.*, **29**, 363, (1974)
4. DEMPSTER, A. J., *Phys. Rev.*, **11**, 316, (1918)
5. BARBER, N. F., *Proc. Leeds Phil. Lit. Soc.*, **2**, 427, (1933)
6. STEPHENS, W. E., *Phys. Rev.*, **45**, 513, (1934)
7. DAVIS, W. D. and VANDERSLICE, T. A., *Trans. 7th Nat. Symp. Vacuum Technol. 1960*, Pergamon Press, 417, (1961)
8. WERNER, H. W., *Vacuum*, **33**, 521, (1983)
9. BLEAKNEY, W. and HIPPLE, J. A., *Phys. Rev.*, **53**, 521, (1938)
10. PERKINS, G. D. and CHARPENTIER, D. E., *Trans. 4th Nat. Symp. on Vacuum Technol. 1957*, Pergamon Press, 125, (1958)
11. KORNELSEN, E. V., *Proc. 19th Phys. Electronics Conf. MIT*, 156, (1959)
12. HUBER, W. K. and TRENDELENBURG, E. A., *Trans. 8th Nat. Vac. Symp. and 2nd Internat. Cong. Vac. Sci. Technol. 1961*, Pergamon Press, 592, (1962)
13. ANDREW, D., *Trans. 3rd Internat. Vac. Cong.*, Pergamon Press, **2**, 527, (1967)
14. STEPHENS, W. E., *Phys. Rev.*, **69**, 691, (1946)
15. WILEY, W. C. and McLAREN, I. H. *Rev. Sci. Instrum.*, **26**, 1150, (1955)
16. POSCHENRIEDER, W. P. and OETJEN, G. H., *J. Vac. Sci. Technol.*, **9**, 212, (1972)
17. SMITH, L. G., *Rev. Sci. Instrum.*, **22**, 115, (1951)
18. SOMMER, H., THOMAS, H. A. and HIPPLE, J. A., *Phys. Rev.*, **82**, 697, (1951)
19. BERRY, C. E., *J. Appl. Phys.*, **25**, 28, (1954)
20. BRUBAKER, W. M. and PERKINS, G. D., *Rev. Sci. Instrum.*, **27**, 720, (1956)
21. ALPERT, D. and BURITZ, R. S., *J. Appl. Phys.*, **25**, 202, (1954)
22. KLOPFER, A. and SCHMIDT, W., *Vacuum*, **10**, 363, (1960)
23. VAN DER WAAL, J., *Suppl. Nuovo Cim.*, **2**, 1, (1963)
24. BENNETT, W. H., *J. Appl. Phys.*, **21**, 143, (1950)
25. SHCHERBAKOVA, M. H., *Zh. Tekn. Fiz.*, **27**, 599, (1957)
26. VARALDI, P. F., SEBESTYEN, L. G. and REIGER, E., *Vakuum Tech.*, **7**, 13; 46, (1958)
27. ROBINSON, N. W., *The Physical Principles of Ultrahigh Vacuum Systems and Equipment*, Chapman & Hall, 118, (1968)
28. SRIDHARAN, R., *Vacuum*, **31**, 159, (1981)
29. TRETNER, W., *Vacuum*, **10**, 31, (1960)
30. PAUL, W. and STEINWEDEL, H., *Z. Naturforsch*, **A8**, 448, (1953)
31. DAYTON, I. E., SHOEMAKER, F. C. and MOZLEY, R. F., *Rev. Sci. Instrum.*, **25**, 485, (1954)
32. PAUL, W., REINHARD, H. P. and VON ZAHN, V., *Z. Physik*, **152**, 143, (1958)
33. DENISON, D. R., *J. Vac. Sci. Technol.*, **8**, 266, (1971)
34. BARIL, M. and SEPTIER, A., *Revue Phys. Appliq.*, **9**, 525, (1974)
35. DAWSON, P. H., *Proc. 7th Internat. Vac. Cong., 1977.* Österreichische Gesellschaft für Vacuumtechnik, Vienna, 173, (1977)
36. BRUBAKER, W. M., *Advances in Mass Spectrometry*, **4**, Elsevier, 293, (1968)
37. BRUBAKER, W. M., *Proc. 6th Internat. Vac. Cong. 1974, Japan J. Appl. Phys. Suppl.*, **2**, 179, (1974)

38. FITE, W. L., *Rev. Sci. Instrum.*, **47**, 326, (1976)
39. BRINKMANN, U., *Int. J. Mass Spectrom. Ion Phys.*, **9**, 161, (1972)
40. HOLME, A. E., *Int. J. Mass Spectrom. Ion Phys.*, **22**, 1, (1976)
41. HOLME, A. E., SAYYID, S. and LECK, J. H., *Int. J. Mass Spectrom. Ion Phys.*, **26**, 191, (1978)
42. DAWSON, P. H. (Editor), *Quadrupole Mass Spectrometry and its Application*, Elsevier, (1976)
43. BARGERY, G. J. and BALL, G. W., *Proc. 4th Internat. Vac. Cong. 1968.* Inst. of Physics, 695, (1968)
44. HOLME, A. E., THATCHER, W. J. and LECK, J. H., *Vacuum*, **22**, 327 (1972)
45. REICH, G., *Proc. 7th Internat. Vac. Cong. 1977.* Österreichische Gesellschaft für Vacuumtechnik, Vienna, 197, (1977)
46. VON ZAHN, U., *Rev. Sci. Instrum.*, **34**, 1, (1963)
47. GOODRICH, G. W. and WILEY, W. C., *Rev. Sci. Instrum.*, **32**, 846, (1961)
48. ESCHARD, G. and MANLEY, B. W., *Acta Electronica*, **14**, 19, (1971)
49. CHISHOLM, T., HUBREGTSE, J., WESTON, G. F. and WINDSOR, E. E., *Residual Gases in Electron Tubes*, Academic Press, 165, (1972)
50. FITE, W. L. and IRVING, P., *J. Vac. Sci. Technol.*, **11**, 315, (1974)
51. BRETH, A., DOBROZEMSKY, R. and KRAUS, B., *Vacuum*, **33**, 73, (1983)

Chapter 6

Ultrahigh vacuum line components

6.1 Scope

In any vacuum system there is usually a requirement for valves, demountable seals, electrical and mechanical feed-throughs, etc. which are classed together here as vacuum line components. Most vacuum equipment manufacturers include a range of such components amongst their products, with compatible interconnections which can be assembled with pumps and gauges to form the complete vacuum system.

For ultrahigh vacuum applications such components must conform to the rigid material requirements previously discussed in Chapter 2 so as to present the minimum source of gas to the vacuum system. In particular, they should be capable of being baked to at least 250°C and give a low outgassing rate and negligible gas permeability when forming part of the chamber walls. In general this precludes synthetic materials such as elastomers, epoxy resins or plastics, which are employed extensively at higher pressures, above 10^{-5} Pa. The elastomers for example, because of their high elasticity, are especially useful as a gasket material in demountable seals and valves where repeated opening and closing is required. Harder materials such as PTFE have properties which make them suitable as seatings or bushes for mechanical movements. Although there are some synthetics which can be baked above 200°C, in general there are no materials with similar mechanical properties which fulfil all the requirements of ultrahigh vacuum applications. The vacuum engineer has to resort to rather different approaches to arrive at equivalent components for pressures below 10^{-6} Pa. The design and development of such components has been an integral part of the ultrahigh vacuum studies since the 1950s and today most manufacturers produce a range of ultrahigh vacuum line components. Although less convenient and more expensive than their higher pressure counterparts, these provide a practical answer to the stringent environmental conditions imposed in attaining ultimate pressures below 10^{-6} Pa.

In this chapter the general design of such components is discussed, with emphasis on the most recent developments. Also included is the liquid nitrogen replenisher, an essential component for systems having cold traps or sorption pumps.

6.2 Demountable seals

Demountable seals can be defined as the static seals made between the various vacuum components which can be broken and resealed to facilitate dismantling for system changes or maintenance. For ultrahigh vacuum application such seals are almost entirely in the form of stainless steel flanges clamped or bolted together with some form of gasket compressed between them. The properties of the gasket are such that it must be deformable under the minimum pressure so that the gasket material can flow and fill surface irregularities on the flanges. On the other hand, it should be resilient enough to maintain contact over a range of compressions to allow for expansion and possibly flexing. Natural or synthetic rubber is an almost ideal gasket material, giving leak-tight joins with low loads on the flange bolts and being resilient enough to be re-usable several times.

Because in general the gasket surface in contact with the vacuum is minimal and the gas diffusion path length through the gasket is relatively long, the demands on degassing rate and gas permeability are less stringent than for the rest of the vacuum envelope. Synthetic rubbers are therefore suitable for vacuum equipment. However, for ultrahigh vacuum systems it is essential that the material used for the gasket will withstand temperatures up to, say, 250°C in order that the flanges can be baked with the rest of the system to reduce the outgassing rates. Few rubber-type materials can be raised to such temperatures without changing their physical properties or decomposing. Nevertheless there are one or two which can withstand such temperatures and have a low enough outgassing rate after baking to be of interest, especially if the pressure requirements are at the upper end of the ultrahigh vacuum range $>10^{-7}$ Pa and if the baking temperature can be lowered slightly. The outgassing rate and gas permeation constant of suitable materials were listed in *Tables 2.7* and *2.8* respectively in Chapter 2, where such materials were discussed.

The best known and most used synthetic material for ultrahigh vacuum application is Viton A, a product of du Pont, which is an elastomer with properties suitable for flange seals. It can be baked up to 200°C and can be employed in standard 'O'-ring seals of the conventional type where the elastomer in the form of a toroid is compressed between a flat flange and one having a groove, usually of trapezoidal cross-section, locating the toroid. Details of such seals are given in most textbooks on vacuum technology, for example that by Pirani and Yarwood[1]. Polyimides, the best known of which is Kapton-H, have also been used for vacuum seals at pressures down to 10^{-7} Pa[2,3]. It has a lower outgassing rate than Viton A after baking but is a much harder material and unsuitable for 'O'-ring type seals. Seals using polyimides have to be carefully designed because of its high expansion coefficient and the fact that it will suffer permanent distortion if compressed more than 20%. One system that appears to be satisfactory uses a thin sheet of Kapton-H compressed between a groove and a knife edge[3]. Because polyimide is, to some extent, hydroscopic, the unbaked degassing rate can be fairly high. Thus, a system using polyimide would have to be baked each time after it has been opened up to air. Two relatively new elastomers have been introduced by du Pont, namely Kalrez and Viton E60C, which, according to Chernatony[4] have rather better outgassing properties. They are now available from a few

vacuum equipment companies in the form of 'O'-rings but they are relatively expensive.

For lower pressures, synthetic gaskets are not suitable and the vacuum engineer must resort to all-metal seals, i.e. the use of metal gaskets. Metal gaskets are available in a number of forms with material ranging from soft metals such as indium, aluminium and gold to harder metals such as silver, copper, monel and even steel. They require considerably greater pressure than elastomers to ensure flow of the gasket metal into the irregularities of the flange faces and they lack the resilience, so that once the seal has been broken the gasket is rarely re-usable. Indeed Buchter[5] has stated that leak-tight metal seals cannot be made unless the metal in the contact area is permanently deformed by compressing beyond the elastic limit into the plastic deformation range.

The sealing mechanism of metal gasket seals has been investigated by several workers, for example Roth[6] who has investigated the effect of temperature cycling and surface roughness. Although the mechanism is reasonably understood there is no really scientific foundation for seal design analysis and most of the numerous designs of ultrahigh vacuum metal seals that have been reported depend on empirical evaluation. In general, the very soft metals, such as indium and gold, are used as unconfined gaskets between two flat flanges so that they are free to deform inwards and outwards along the flange surfaces. The harder metals are usually partly confined by contoured flanges so that the gasket is only free to deform in one direction.

The unconfined gasket normally consists of a wire ring fabricated by fusing the wire ends together. It is then mounted between two stainless steel flanges which are machined accurately flat and to a fine finish. Care must be taken to avoid any radial scratches on the flange surfaces. The ring may be held in position with a spider as in *Figure 6.1*, or just positioned with a suitable tool whilst the bolts are tightened. The main criterion for the design is that the wire should be of fairly uniform thickness. The diameter must be large enough to form a reasonable contact area along the flange surface when compressed but not so large that there is a danger that the atmospheric pressure will push it

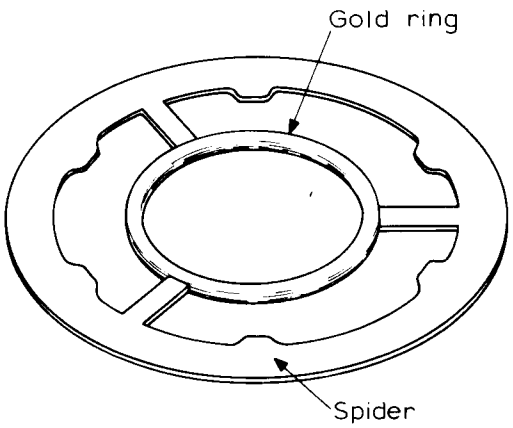


Figure 6.1 Gold wire gasket with positioning spider

inward and burst through. 1.0 mm diameter wire is about the maximum size, but 0.5 mm diameter is normally employed compressed to about 0.2 mm, i.e. the spider in *Figure 6.1* is 0.2 mm thick.

An alternative arrangement is offered by Leybold–Heraeus who market an aluminium disc with a diamond-shaped raised sealing rim. The rim acts in the same way as the wire whilst the disc locates the sealing rim and limits the compression. The advantage of the ring type of seal is that the flanges are fairly easily machined and the wire rings can be made up by the user. The disadvantage is that in making the seal and in subsequent temperature cycling the wire tends to weld to the flanges which may necessitate resurfacing the flanges before they can be re-used. They are therefore most suitable for semipermanent seals where long periods elapse before they are required to be opened. Another problem is that the flange surfaces are vulnerable to damage when exposed. Quite small scratches can degrade the vacuum performance.

Gold is the most satisfactory metal for such seals. It is not affected by contaminants such as oil vapours and it can be baked up to 500°C. Also the value of the gold can be reclaimed. Gold wire seals are sold commercially, particularly for large diameter flanges.

For smaller flanges the trend has been to move to copper gasket seals using designs which have become standardized. Copper, being a harder metal than gold, requires a larger pressure to cause deformation. It is unsatisfactory as a wire seal and although 'diamond section' ring gaskets can be used, most seals employ relatively 'heavy section' gaskets which are partially confined.

A number of configurations for copper gasket seals have been proposed in the literature and a selection are shown in cross-section in *Figure 6.2*. The designs are self-explanatory and generally rely on a sharp-edge biting into the gasket. This gives the maximum specific force for deformation of the copper

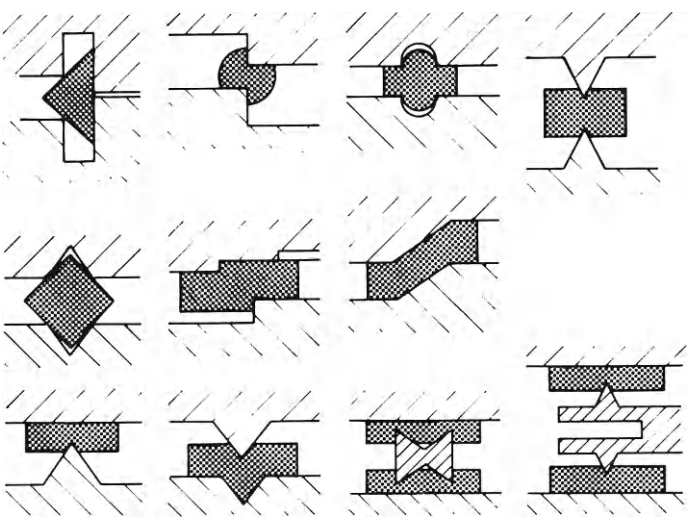


Figure 6.2 Some copper gasket seal configurations (according to Buchter[5], Copyright © 1979, John Wiley & Sons Inc. Reprinted with permission)

from the minimum force on the flange bolts. The relative merits of the various designs are difficult to assess and in the early days of ultrahigh vacuum several of these designs were exploited commercially. The problem was that the different flange designs were not compatible and equipment from different manufacturers had to be modified to allow interconnection.

A similar problem had arisen for the elastomer 'O'-ring seals. In the early 1960s the International Standards Organization (ISO) in co-ordination with Pneurop (an assembly of European manufacturers of compressors, vacuum pumps and pneumatic tools) set about drawing up international standards for vacuum flanges as part of a wider study of vacuum equipment. In the case of the UHV metal gasket seals, recommended specifications have been arrived at for a range of flanges to cover the interconnection of different size tubing. The specification details the flange diameter, number of bolt holes and their position, etc. but does not specify the actual sealing mechanism. *Table 6.1* gives the dimensions of those in common usage. Although the sealing method was not specified, it was clear that, since contoured flanges were required, the sealing configuration had also to be agreed. Most manufacturers have opted for a Varian design[7] marketed under the trade name of ConFlat® and this has become the accepted standard.

TABLE 6.1. ISO metal flange recommended dimensions

Nominal tube i.d. (mm)	*Outside diameter of flange* (mm)	*Diameter of bolt hole centres* (mm)	*Number bolt holes*	*Diameter of bolt holes* (mm)
16 ($\frac{3}{4}''$)	34	27	6	4.3
35 ($1\frac{1}{2}''$)	69.5	58.7	6	6.6
63 ($2\frac{1}{2}''$)	113.5	92.1	8	8.4
100 (4″)	152	130.2	16	8.4
150 (6″)	202.5	181	20	8.4
200 (8″)	253	231.8	24	8.4
250 (10″)	306	283.5	32	8.4

The ConFlat flange seal is illustrated in *Figure 6.3*. It consists of two symmetrical flanges each having an outside rim to confine the flat copper gasket and a concentric knife edge of a particular cross-section. When bolted together the knife edges bite into the copper gasket causing lateral flow. The outer rim of the flange, however, restricts the flow, increasing the interface pressure to ensure that the copper fills any imperfections in the flange surface. Properly employed this arrangement gives virtually 'leak-tight' seals, in that the leak rate is undetectable with a helium leak detector (better than 10^{-12} Pa $m^3 s^{-1}$). Another advantage of the ConFlat flange is that the design allows the knife edge to be an integral part of the component whilst the outside ring with the bolt holes can be free to rotate, an important consideration in suitably positioning components such as valves. The outside rim also protects the knife edge from mechanical damage. The flanges are bakable up to 450°C. If long periods of baking or baking in an oxygen environment are required, the copper gasket can be silver plated to prevent excessive oxidation. Because of the plastic deformation the gasket is not re-usable.

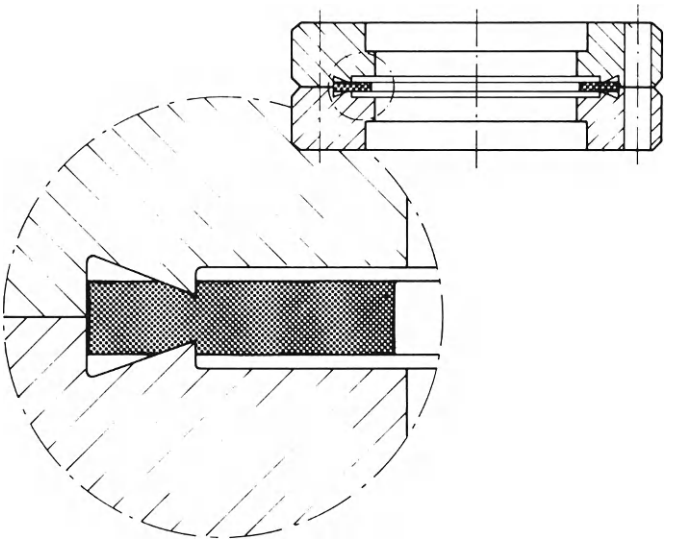


Figure 6.3 Basic design of the ConFlat® flange seal from Varian Associates

Metals such as monel and stainless steel can also be used as gaskets by exploiting their springy characteristics. Various shaped cross-section gaskets, such as X- and V-shaped, used in other applications have been adapted for vacuum use. Of particular interest is the use of metal tubular 'O'-rings. 'O'-rings of stainless steel tubing are manufactured for seals and can be used in similar systems to the elastomer 'O'-ring. The tubes may be filled to atmospheric pressure or may be pressurized. For ultrahigh vacuum, the seal is not sufficiently leak tight using the steel ring alone. However, it can be coated with a softer metal, for example copper, or with Teflon. Teflon can be baked to 250°C and has a low outgassing rate, although its gas permeability is relatively high. Teflon coated rings have been used at the Cavendish Laboratory for vacuum systems pumping down to 5×10^{-8} Pa[8]. A variation on this system has been described[9] in which the 'O'-ring consists of a helical spring core with a soft metal sheet wrapped around it. There is also a commercial design using a ring of C-shaped cross-section, *Figure 6.4*. Although they are comparable in price to the copper gasket they are less critical on the flange contour and do not need such a high compression force. They can be re-used provided they have not been distorted or damaged by the previous sealing and/or temperature cycling.

Whatever metal gasket sealing mechanism is employed, there are certain essential design features common to them all that should be observed for ultrahigh vacuum application. First, the material used must be of UHV quality. The stainless steel must not be porous or contain inclusions likely to affect the knife edge or surface finish. The gasket should be chemically pure and in the case of copper, OFHC (oxygen free high conductivity) copper should be used. Secondly, the surfaces should be free of scratches and be clean. In particular, they should be free from dust and grease and should not be handled after cleaning without gloves. Thirdly, when the flanges are connected to the

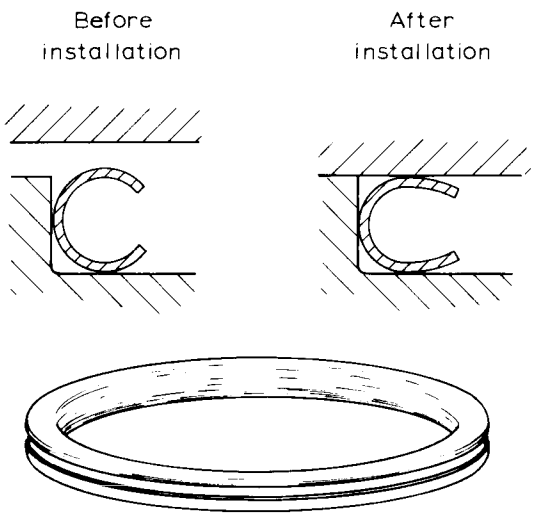


Figure 6.4 C-shaped ring gasket from Avica Equipment Ltd.

tubing they should be welded on the inside with an argon arc to ensure there are no voids which could represent a virtual leak. Care must be taken not to distort the flange during welding; a single pass with a penetration depth of about 1.5 mm is suitable. The flanges should be bolted together with high tensile strength stainless steel bolts with a thermal expansion matched to that of the flanges. To ensure even pressure on making the seal, the bolts should be tightened in a suitable sequence. The actual torque required need not be measured as the flanges are normally tightened until either they touch or a fixed gap, measured with a feeler gauge, is reached. Finally the gaskets should not be re-used and it is false economy to attempt to do so. The exception is the hard springy metal seal, such as the 'O'-rings of stainless steel, which can be used more than once in some circumstances. Also, Balzers have recently announced a re-usable aluminium seal which can be used in ISO clamped-flange designs. The seal consists of two metal gaskets and an X-shaped cross-sectioned support ring placed between them. On sealing, the gaskets are forced into the cavity of the X cross-section. To re-use, the gaskets are turned over and again pressed into the grooves. It can only be re-used if low baking temperatures, $<150°C$, are employed.

6.3 Mechanical feed-throughs and dynamic seals

In many vacuum systems there is a need to transmit mechanical movement through the vacuum envelope. This may be to position a component or to provide a continuously moving part within the vacuum. In this section the various methods of achieving mechanical transmission whilst maintaining ultrahigh vacuum conditions are considered. The applications can be conveniently divided into two categories; those in which manipulation of components over relatively small distances or rotation at low speeds, up to, say,

1000 rev/min are required and those requiring large movements or rotation at high speed and/or with a high torque.

For vacuum equipment operating at pressures above 10^{-4} Pa, the first category is served by 'O'-ring seals or grease-packed lip washer seals using vacuum grease or oil as the lubricant. Although the leak rate for such seals can be low, they are not bakable and therefore unsuitable for ultrahigh vacuum. Such seals can be improved by using Viton A 'O'-rings lubricated with graphite, but in general these are only satisfactory at the upper end of the ultrahigh vacuum pressure range. Fortunately there is a satisfactory solution which meets the stringent vacuum requirement and which is not excessively expensive. The method uses flexible bellows as part of the vacuum envelope, to manipulate the component from outside. This is illustrated in *Figure 6.5* which shows simple 'wobble sticks' from Vacuum Generators Ltd.

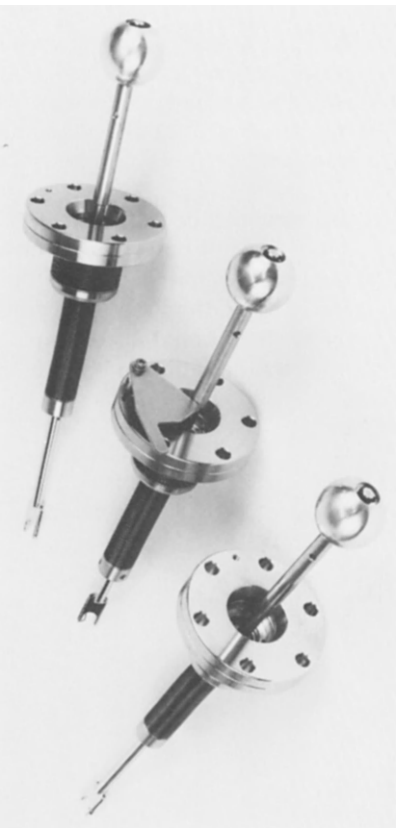


Figure 6.5 'Wobble sticks' using bellows. (Courtesy Vacuum Generators Ltd.)

Two types of bellows are used for such manipulators, hydraulically formed bellows and edge welded bellows, illustrated in *Figure 6.6*. The formed bellows, as the name implies, are formed from tubing by forcing the convolutions into a collapsible die using hydraulic pressure. The edge welded bellows consist of preformed diaphragms, argon arc welded together along alternate inner and outer rims. Various profiles can be used for the diaphragms. The formed

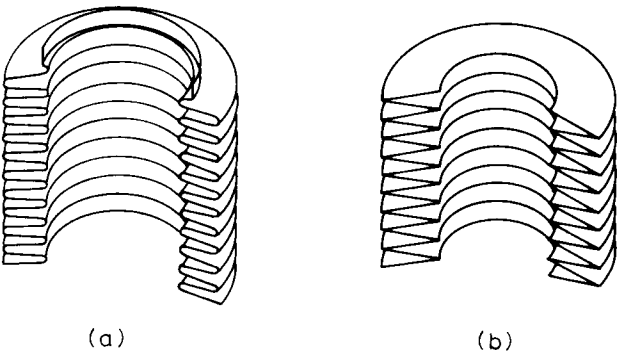


Figure 6.6 Types of metal bellows. (a) Hydraulically formed; (b) Edge welded diaphragms

bellows are considerably cheaper than the welded bellows but the latter give a superior performance in terms of flexibility and in particular a greater linear motion along the axis per unit length of bellows. The vacuum performances are similar. Both can be made of 18/8 stainless steel and both can be baked up to 450°C. The bellows are used under compression and preferably should not be extended much beyond their free length.

The manner in which linear motion in any direction can be transmitted via the bellows is fairly obvious from *Figure 6.5*. What may not be so clear is how continuous rotation of a component may be achieved. *Figure 6.7* shows one possible arrangement using ball and socket type joints. Other configurations have been employed based on a similar principle. Rotational speeds up to 2000 rev/min have been achieved, but mainly they are designed to operate at much lower speeds or under manual control.

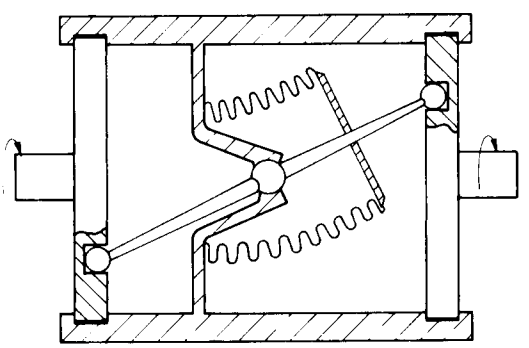


Figure 6.7 Schematic diagram of a rotating drive with metal bellows

Linear and rotational motion feedthroughs based on the bellows system are marketed by most vacuum components manufacturers. They vary from simple 'wobble sticks' to precision manipulators with vernier adjustments and are usually fitted with flanges for connection on to the vacuum system. An example of a more complex manipulator is shown in *Figure 6.8*.

For the second category of applications requiring greater movement or

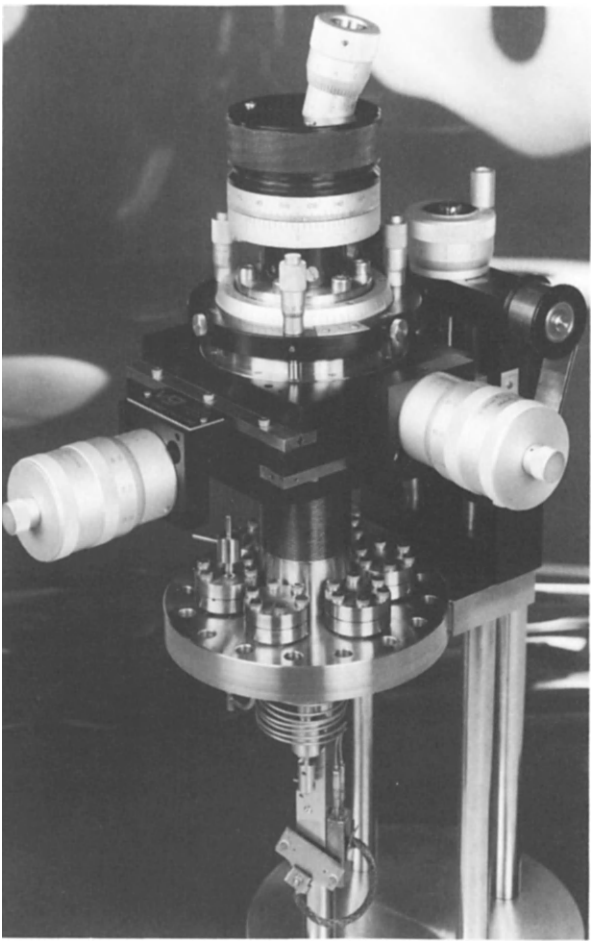


Figure 6.8 Ultrahigh vacuum precision manipulator. (Courtesy Vacuum Generators Ltd.)

faster rotational speeds the bellows system is less satisfactory. One convenient method, which also obviates the need for the shaft to pass through the vacuum envelope, is to use magnetic or electromagnetic coupling. For example the coil of an electric motor can be placed outside the vacuum envelope whilst the rotor is mounted inside to obtain rotary motion. Since the rotor must not be a source of gas it should be of minimum bulk and of a suitable material which can be baked. A wound-coil rotor, unless encased in a sealed envelope, would be unsatisfactory and therefore a.c. induction motors with cage-type rotors or reluctance motors with magnetic rotors are employed. The problem with such arrangements is that the air gap between rotor and stator has to be sufficient to accommodate the vacuum envelope. This results in a reduction in the torque and in the efficiency. In general such motors are restricted to low torque requirements such as rotating small piece parts, although high speeds can be obtained. An alternative method is to produce a rotating magnetic field by rotating magnets outside the vacuum. This type of magnetic drive allows a

larger air gap between the magnetic drive and magnetic rotor and thus a thicker vacuum envelope. Several designs were used in the 1960s, for example Coenraads and Lavelle[10], and more recently Budgen[11] described the drive for a mechanical booster pump using an arrangement with three rotating magnets.

The movement of a magnet or ferromagnetic component inside the vacuum by a magnet outside is also an obvious method of obtaining extensive linear movement of components. The external magnet can either be moved over the required distance in parallel with the internal component or it can be rotated to drive a leadscrew. The only requirement here is that the vacuum envelope should not interfere with the magnetic coupling.

Thus, for the majority of applications, mechanical movements within an ultrahigh vacuum system can be effected by one of the above methods, without the need for the drive shaft to pass directly through the vacuum envelope. There are, however, some applications where a high torque is required at relatively high speeds and the vacuum engineer needs to resort to a dynamic seal on the drive shaft at the vacuum/atmospheric interface. Normal dynamic seals used on rotating shafts exploit mating surfaces with a thin oil film between them and are similar to bush bearings. Using low vapour pressure grease or oil, such seals can be used down to pressures of the order of 10^{-4} Pa. Oil can be eliminated by using PTFE bushes which can then be baked up to 200°C. A suitable shaft seal could be the BAL-seal manufactured by the BAL-seal Engineering Corp. The seal is a U-shaped PTFE ring with the opening directed towards the atmospheric pressure side with a special spring imbedded inside the U-cavity as illustrated in *Figure 6.9*. It is applied like an 'O'-ring and is suitable for rotating or reciprocating shafts. By using two seals and

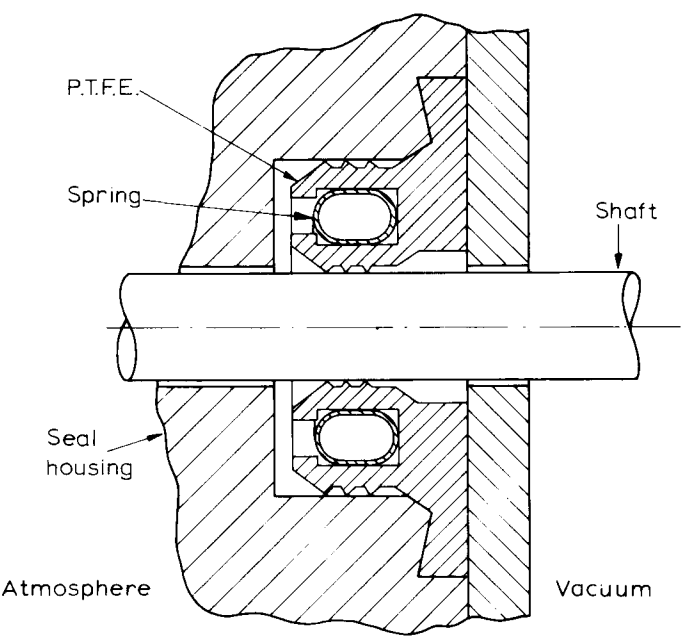


Figure 6.9 Basic design of PTFE dynamic seal from BAL-seal Engineering Corp.

evacuating the space in between with a backing pump, the leak rate is minimal. Similar PTFE dynamic seals are available from vacuum component manufacturers fitted with suitable vacuum flanges.

A new type of seal has been introduced recently for high vacuum applications (pressures down to 10^{-6} Pa) which is becoming more popular. It is based on the use of a magnetic fluid and is marketed by the Ferrofluidic Corporation[12]. The magnetic fluid consists of submicroscopic magnetic particles colloidally suspended in a carrier liquid which, for vacuum application, is a low vapour pressure oil. In the absence of a magnetic field the liquid behaves in its normal fashion. On application of the magnetic field the motion of the suspended particles along the applied field transports the oil by osmotic forces, causing the suspension as a whole to move. The liquid properties are retained but the liquid can now be contained within the magnetic field to build up an oil barrier which will stand up to a differential pressure. An illustration of such a seal for a rotating shaft is given in *Figure 6.10*. It uses a permanent

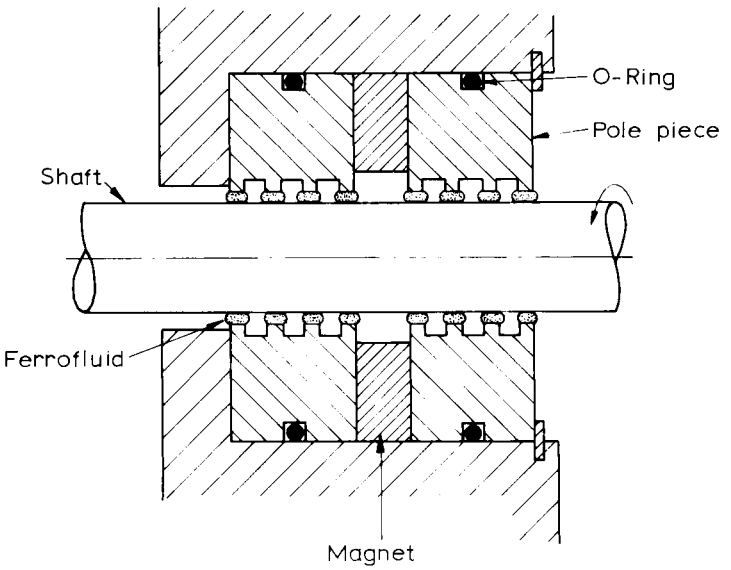


Figure 6.10 Ferrofluidic rotating shaft seal from Ferrofluidics Corp.

magnet and multistage seal to stand up to the atmospheric pressure difference. Since there is little or no friction the seal has a long life and allows a high rotational speed of the shaft; many are rated above 5000 rev/min and special designs have performed above 50 000 rev/min. They are relatively cheap compared with alternative systems and are conveniently designed with standard vacuum flange connections. Unfortunately there are no dynamic seals known to the author which are completely satisfactory for pressures below 10^{-7} Pa.

Whatever drive mechanism is used there will be a need to lubricate the moving components which are within the vacuum envelope, otherwise binding and even cold welding can occur. This poses the problem of finding a suitable lubricant which will not constitute a contamination and which can withstand

the baking temperature, criteria which rule out most oils and greases. It is a problem which is also significant in the space programme and as a consequence has been extensively studied. Solutions vary from soft metal coatings on ball bearings such as lead and silver, to dry films which combine molybdenum disulphide with graphite. PTFE bushes or coatings are also very suitable if the baking temperature can be limited. Magnetic bearings which hold the component in levitation have also been described. Thin film coatings of either soft metals or of hard low-friction materials, such as molybdenum-sulphide and glassy carbon, deposited on to the bearing surfaces probably offer the best solution for bearings in systems where ultrahigh vacuum pressures are to be attained.

6.4 Valves

Valves are probably the most important vacuum line component, allowing isolation of the various regions of the vacuum system essential for pumping down, processing and maintenance. Because of the number of roles they play in the vacuum system there is a wide variety of designs from the small air inlet valves to the large open-conductance gate-valves used, for example, in the beam line of particle accelerators. There are, however, some basic design principles on which most of the commercial valves depend and by describing examples, the overall picture of ultrahigh vacuum valve design for the range can be presented. The basic designs depend on the application. They depend on the required open conductance or tube bore in which the valve is to be connected, whether the valve is to be an on–off device or used to control the gas flow rate and how stringent the vacuum conditions have to be. For some applications a straight through path may be required to pass a charged particle or radiation beam in the open position. The two main design types are the right-angle valve so called because of the position of the vacuum tube connections and the straight through valve where the tubes are in line. In the latter case, a common design is the gate valve which facilitates a large unimpeded conductance.

There are certain criteria that all ultrahigh vacuum valves have in common. They should have a minimum leak rate in the closed position (e.g. $<10^{-11}$ Pa $m^3 s^{-1}$), have a maximum conductance in the open position and not be a source in themselves of gas contamination. The last criterion implies using materials compatible with the low pressure requirements with negligible gas permeability and which can be outgassed by baking to a temperature of at least 200°C but preferably to 450°C. The leak rate and baking criteria are the same as those applying to demountable seals and it is not surprising therefore that many of the valve seating designs resemble the configurations used in the flange seals. Similarly, operating the opening and closing mechanism of the valves requires a mechanical feed-through to the same specification as those described in the previous section.

At the upper end of the pressure range, $>10^{-7}$ Pa, elastomer sealing techniques can be adapted for valve application. It is essential that the synthetic material used can be baked up to, say, 200°C, for example Viton A, and that the mechanical feed-through is also compatible with the vacuum requirements. A number of designs for this pressure region have been

proposed and several alternative designs are sold on the market. For relatively small-bore tubing, right-angle valves are normally employed, but this is for convenience in the design rather than any vacuum consideration. Probably the commonest design for this type of valve exploits the ubiquitous 'O'-ring seal; an example is illustrated in *Figure 6.11*. The body of the valve is stainless steel and the drive mechanism is operated through a bellows seal. Using Viton A 'O'-rings the valve is bakable up to 200°C. Valves of this type are normally used for isolation in vacuum lines of 13–35 mm diameter but they can be obtained for connecting to tubes of up to 150 mm diameter. For the large diameter bores, however, the gate valve offers advantages. It takes up less vacuum line space and allows a straight-through path.

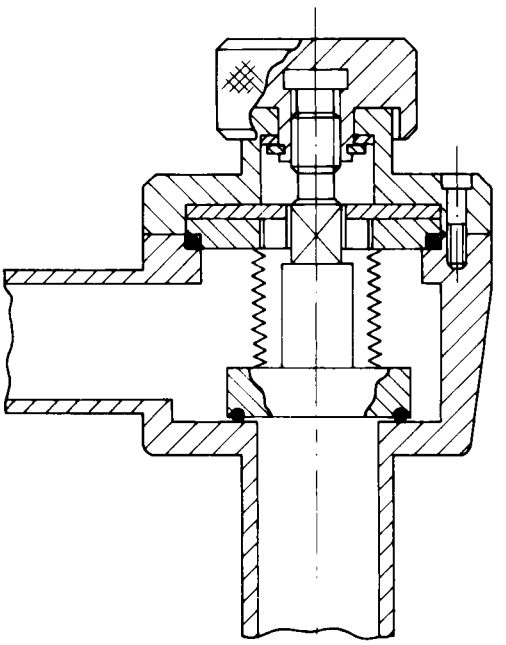


Figure 6.11 'O'-ring elastomer sealed right-angle valve from Leybold–Heraeus

The basic design of a gate valve is shown diagrammatically in *Figure 6.12*. The shaft first moves the seal plate into the gate and then applies the transverse motion to close the valve by an 'over-shoot' mechanism, shown here as ball bearings being pushed out of indentations, a method exploited by VAT who specialize in such valves. In this example the transverse force is applied between the seal plate and a backing plate, but in some designs the force is applied between the seal plate and the sliding carriage. The seal is made with a Viton A 'O'-ring and for the bakable version a bellows seal is employed on the shaft.

Valves required to control the flow of gas in this pressure region are usually based on the needle-valve principle. The diameter and taper of the needle

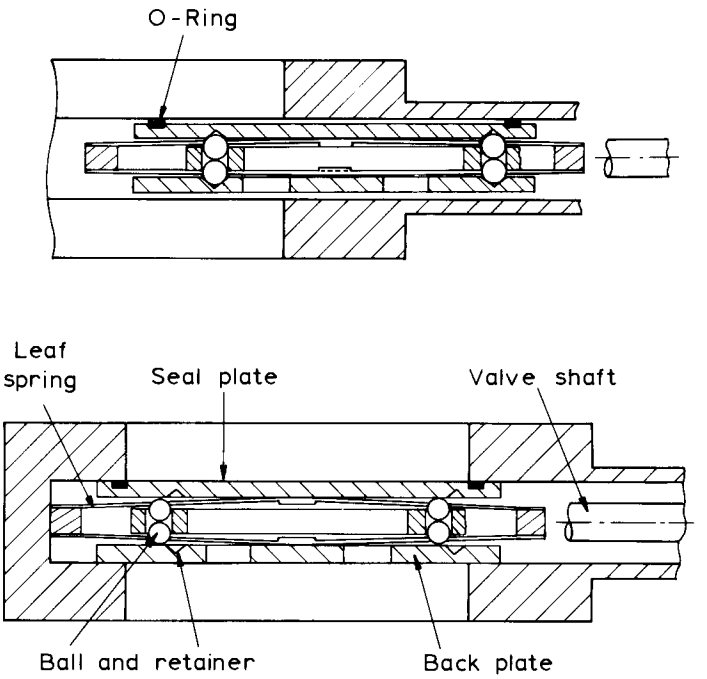


Figure 6.12 Elastomer sealed gate valve, basic design patented by VAT (Vakuum-Aparate-Technik, Haag, Switzerland)

determines the flow rate range. The needle and seating are of hard metal, often stainless steel, and in some designs a soft metal is deposited on the seating to give a better vacuum seal when closed. Some are provided with a separate seal in the closed position using a Viton 'O'-ring. Few of the designs can be baked, although they may be fitted with a bellows seal on the drive shaft and use materials compatible with high vacuum requirements. The control is normally via a micrometer type head acting on a differential screw which can be calibrated.

For lower pressures, below 10^{-6} Pa, 'all-metal' valves are required. The first all-metal valve was developed by Alpert[13] at Westinghouse at the inception of ultrahigh vacuum techniques in the early 1950s. It was constructed in two parts, the main valve body and the drive mechanism which was detached for baking. A schematic diagram of the design is given in *Figure 6.13*. The valve body consisted of a copper cup with 2 orifices into which the copper tubes of approximately 6 mm bore were sealed. The front of the cup was sealed with a Kovar diaphragm. A nose was mounted at the centre of the diaphragm and had a polished 45° cone which came into contact with the edge of the central orifice when the valve was closed. The diaphragm, nose and tubes were all brazed on to the copper cup in one operation in a hydrogen furnace to prevent any oxidation. When the valve is first closed the copper flows sufficiently to give a reasonable surface area in contact with the nose cone. A force of about 10 kN is required. The driving mechanism is a differential screw which allows precise positioning of the cone relative to the seating. Because of this the valve can be used as a gas flow control valve. The diaphragm allows a cone

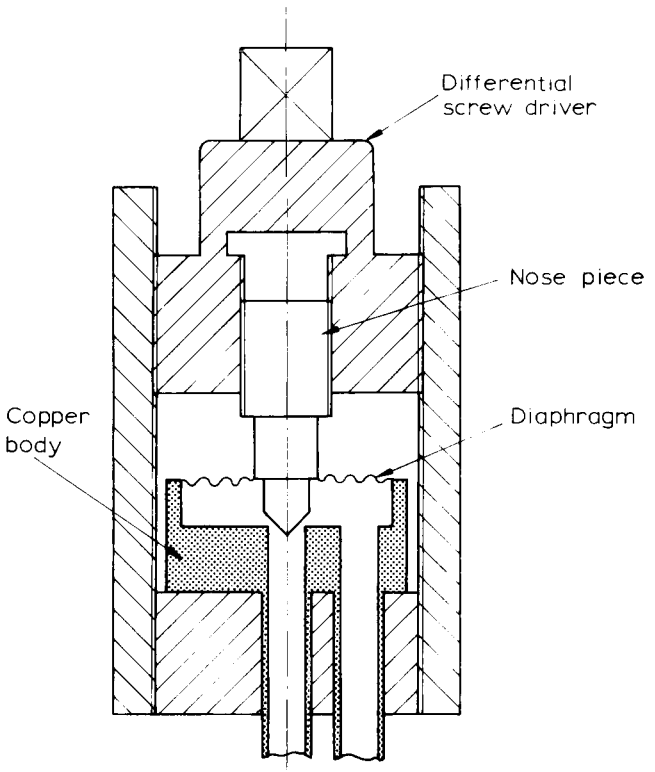


Figure 6.13 Schematic diagram of the Alpert valve

movement of about 3 mm. The design or modifications of it were taken up by several companies and the same basic design for connecting to glass systems is available today, although inconel may be used instead of the copper with a stainless steel diaphragm. Such valves were bakable up to 450°C in the open position whilst the later designs can also be baked closed.

Such valves were limited in size, giving an open conductance of around 5×10^{-4} m^3s^{-1}. For larger valves, bellows are used instead of the diaphragm to give a larger movement and therefore a higher open conductance. The cone and orifice system was found unsuitable for larger diameters and the popular choice has been the knife edge seal. Most manufacturers offer such valves in standard sizes to connect to ISO flanges, i.e. nominal vacuum tube diameters of 16, 35, 63 mm, etc. (see *Table 6.1*). The knife edge is normally part of the main body of the valve and made of stainless steel. The valve plate has a flat surface which presses on to the knife edge. Since the knife edge will normally cause an indentation on the plate it is essential that the actuating mechanism precisely positions the plate to ensure that the indentation is aligned with the knife edge at each closure. Two types of valve plate are common; in one a copper pad is fitted, whilst in the other a hard material such as stainless steel or sapphire is used. In the latter case one of the surfaces, the knife edge or the plate is usually plated with a thin layer of gold. A copper pad design is illustrated in *Figure 6.14*. Most of the designs have a metal bonnet seal which allows the sealing pad

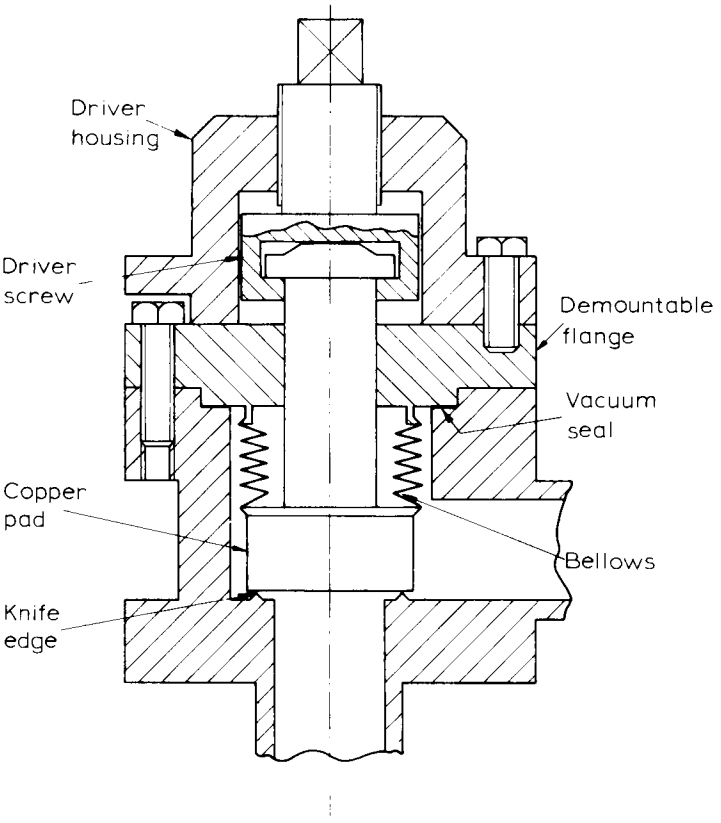


Figure 6.14 Schematic diagram of a knife edge sealed right-angle valve

on the valve plate to be replaced when worn. Some have a seal arrangement which also allows replacement or remachining of the knife-edge.

The pros and cons of using copper pads or a hard seat have been expressed in the literature and in data sheets, but in practice there is little to choose between them. Both can be baked in the open position to 450°C. Most of the designs can also be baked in the closed position, some up to 450°C. In general the hard plate type can be baked to the highest temperature in the closed position whilst rather lower temperatures 300–400°C are recommended for the copper pad designs. This is mainly due to the fact that the copper pad tends to soften by annealing. If there is oxygen likely to be around whilst baking, a situation that would arise with an air-inlet valve for example, the copper pad could become oxidized. However, a high-temperature copper alloy with similar properties can be used as an alternative. The closing torque required is high, particularly for the hard seal version, several Nm, and there can be difficulties in lubricating the actuating mechanism to prevent binding when high temperature cycling is involved. When the valve is to be baked in the open position it is often expedient to remove the actuating mechanism during baking and some designs are constructed to facilitate this.

Modification of the design as shown in *Figure 6.15* will allow the same

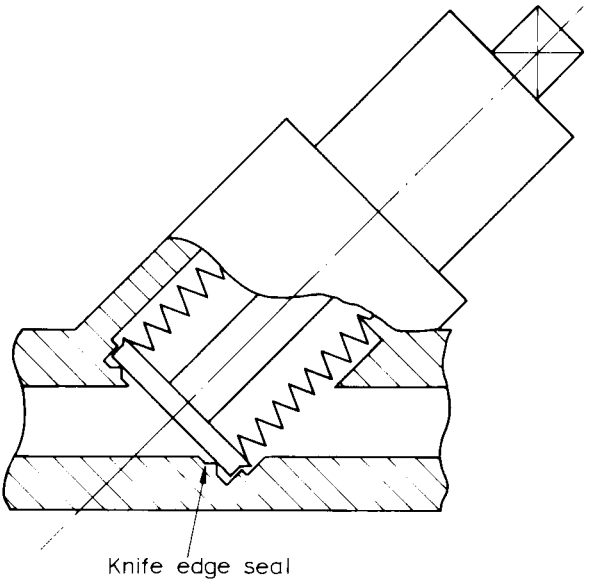


Figure 6.15 Straight-through valve using a knife edge seal

sealing mechanism to be adapted for a straight-through valve. The orifice in the open position will be rather limited, however, and for a large unimpeded path through the valve, a gate valve or similar design of valve is required. The design of an all-metal gate valve involves special technical problems because of the force required to make the gas-tight seal in the closed position, which is at least 10 times that required for an 'O'-ring seal. There is also the need to attain exact positioning of the seal plate relative to the seating for repeated operation. Because of these factors, the designs of all metal gate valves have had varying degrees of success. It has been found to be very difficult to obtain operational lives comparable to those attained with the right-angle valves, especially where frequent temperature cycling is imposed. The larger the valve, the more difficult it is to fulfil the requirements. Most of the large valves, with ports of 150 mm or more diameter, have been commissioned for specific particle accelerators and although they are available on the market they are very expensive, several thousand pounds.

As with the right-angle valve the main seal is commonly the knife edge. If it is used with a copper pad then there is the problem of ensuring alignment of the knife edge with the indentation that is impressed on the copper. The hard metal seal with a thin plating of gold or silver is better in this respect. If the plating is applied to the knife edge, positioning is non-critical. On the other hand a larger sealing force may be required. The actuating mechanism can be similar to that used with the 'O'-ring design. However because of the large force required to close the valve, there is difficulty in providing a mechanical drive with a low enough friction level, bearing in mind that the use of grease or oil lubricants would be unacceptable. In some designs this problem has been overcome by providing a separate, pneumatically operated sealing system. An example of this is a design by Granville–Phillips Corp., where the seal plate is relatively thin and easily positioned and is pressed into position by an annular

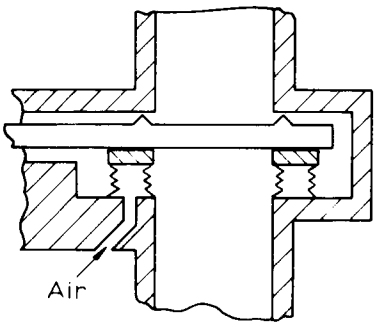


Figure 6.16 Basic principle of gate valve with pneumatic sealing mechanism from Granville Phillips Corp.

bellows welded to the port opposite the seat; the basic principle is illustrated in *Figure 6.16*. A more complex arrangement designed by CERN[14] involved a flat hard seal at each port and a bellows between the two seal plates with softer metal knife edges. The whole system of seal plates and pneumatically operated bellows was swung into position on a short shaft in a 'pendulum' motion and the space between the two seals was separately evacuated to give a low leak rate. An added advantage of the pneumatically operated seal is that it maintains a constant sealing force during the life of the valve.

A novel gate valve design from Varian Associates[15] is claimed to overcome most of the problems. The design is illustrated in *Figure 6.17*. The seal plate is a conical shape of thin metal which is deformed under compression to form a cylindrical seal around its periphery (see inset). The sealing ridge is gold plated. The gate is swung into position in a pendulum motion and the pressure is

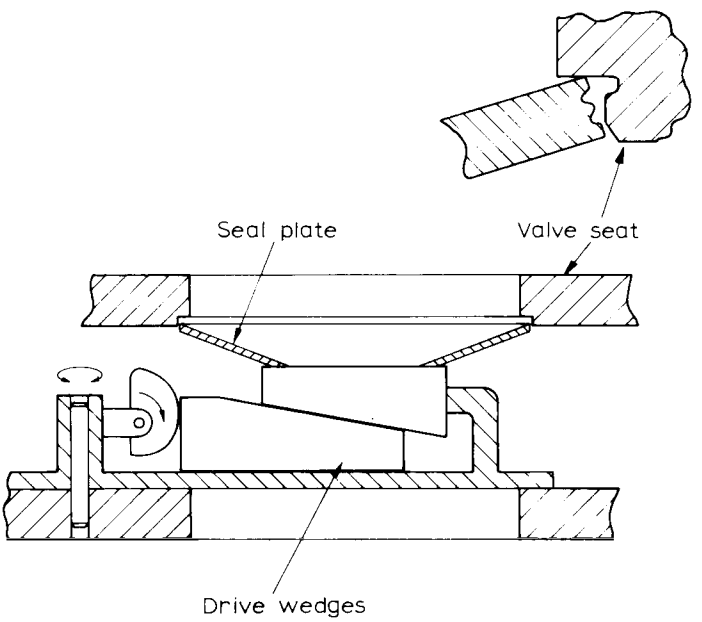


Figure 6.17 Gate valve design from Varian Associates

applied by driving a wedge into the carriage mechanism. The shape of the cone and the fact that the seal is round the edge makes it less sensitive to positioning than the knife edge seal. It is also much lighter in weight.

In general the baking temperatures for gate valves are more limited than for the right-angle valves and most manufacturers do not recommend baking above 300°C.

Because of the precision of alignment between the valve plate and its seating and the exact point of closure, almost any of the all-metal right-angle valves can be adapted with a fine screw adjustment on the actuator to function as a gas flow control valve. Since a fairly restricted movement is all that is required, the use of a diaphragm rather than bellows on the drive mechanism offers the minimum dead volume. To give a fine adjustment the control screw is often applied via a lever. Accurate and reproducible control is claimed for such valves, although few manufacturers produce any figures for the tolerances. For most applications the requirements are defined by the maintenance of a constant pressure drop across the valve and for this purpose most of the control valves are adequate.

6.5 Other vacuum line components

Apart from the need for pipeline components such as flexible connections, T-junctions, side tubes to vacuum chambers, etc., most ultrahigh vacuum systems will also require electrical feed-throughs and viewing ports. There may also be a need for a liquid feed-through which will carry cooling fluid for example.

For the pipeline components, the main criterion is that the material, in most cases an austenitic stainless steel, is of high quality free from inclusions or cavities which could give rise to gas permeation or affect the outgassing rate. The welding and brazing of such components should be made as far as possible along the inside surfaces to ensure that no gas is trapped in the join which could leak into the vacuum system. The joining techniques are discussed in some detail in Chapter 2, Section 2.6.3.

For electrical feed-throughs and viewing ports the controlling design factor is the sealing bond between the ceramic or glass and the metal, which has to withstand temperature cycling during manufacture and use. Because of the brittleness of glass and to some extent also of ceramics, generally such seals can only be made if the temperature coefficient of expansion of the insulator and of the metal are closely matched. This more or less precludes direct sealing to stainless steel where the expansion coefficient is around $17 \times 10^{-6}\,°C^{-1}$ as compared to the values for glass and ceramics which vary from around 3 to $9 \times 10^{-6}\,°C^{-1}$. Fortunately the need in the electronics industry for glass and ceramic to metal seals for thermionic valves has prompted the development of a number of metal alloys which closely match the expansion coefficient of certain glasses and ceramics. This has been discussed in Sections 2.6.2 and 2.6.4. Mainly a NiFeCo alloy in the ratio of 29 : 54 : 17, originally known under the trade name of Kovar, is used. It matches the harder borosilicate glasses and also the high strength alumina ceramics. The glass-to-metal seals are made by directly heating the glass and metal in contact to well above the glass softening point. Ceramic seals are usually made by first metallizing the ceramic by a

sintering and plating technique and then brazing it to the metal. The metal alloys used can then be directly welded to the stainless steel provided the design allows the metal to take up the strain introduced by their differences in expansion coefficients. Sometimes unmatched seals are employed whereby a ductile metal such as indium or gold is introduced between the surfaces to be sealed to take up the thermal expansion differences. The bond is made under pressure at a temperature below the melting point. The method is useful for sealing special windows such as sapphire or quartz although they can also be sealed directly to glass provided a glass is chosen with a matching expansion coefficient. Further details of window seals are to be found in Section 2.6.5.

The high density alumina ceramic is strong enough to withstand the strain of sealing direct to unmatched metals provided the metal can flex sufficiently to compensate for the expansion differences. Thus copper rods can be brazed into alumina ceramics for current leads and direct sealing to stainless steel can be effected if the metal is thin enough at the braze. The techniques required for sealing lead-throughs and windows are very specialized and require considerable skill to perfect. It is therefore fortunate for the vacuum engineer that there is a wide range of electrical feed-throughs available from vacuum component manufacturers which will meet the majority of applications. *Figure 6.18* shows some examples of such components.

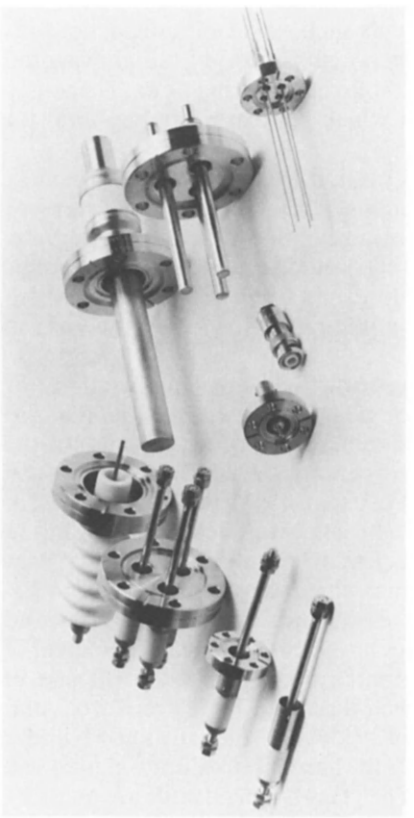


Figure 6.18 Examples of electrical feedthroughs. (Courtesy Vacuum Generators Ltd.)

6.6 Liquid nitrogen replenisher

Although not a vacuum line component, the liquid nitrogen replenisher is nevertheless an almost essential accessory for ultrahigh vacuum systems where either a cooled trap or a sorption pump is employed. If the liquid level in either of these devices falls below a minimum value, gas can be released with disastrous results. Therefore, it is important that the relevant liquid nitrogen container is kept topped up from a reservoir to avoid such an eventuality, particularly if the vacuum system has to be left unattended.

Ideally the replenisher should be fully automatic and for this a sensor is required to assess the liquid level and to operate the liquid transfer mechanism when the level falls below a predetermined value. The transfer system is common to most commercial designs of replenisher and harnesses the pressure build up above the liquid nitrogen in the reservoir to force the liquid across. The basic system is illustrated in *Figure 6.19*.

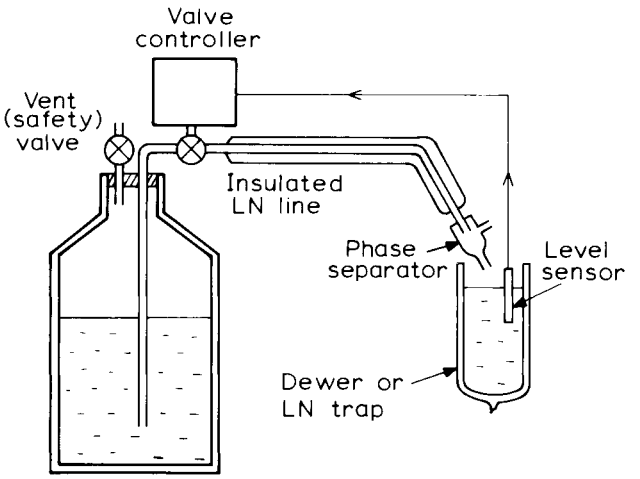


Figure 6.19 Basic arrangement of an automatic liquid nitrogen replenisher

With the container full, the transfer valve and the vent valve are closed. When the level falls to the predetermined value the sensor activates the opening of the transfer valve. The loss rate of liquid nitrogen in the reservoir is normally sufficient to self-pressurize the reservoir. In some systems compressed air is applied to the reservoir to ensure adequate pressure.

Several different sensors have been reported in the literature and are used in commercial designs. One type employs a float on the liquid nitrogen, thus detecting the level mechanically[16]. When the level falls below the predetermined value, a contact attached to the float activates a relay which opens the transfer valve. The problem with this type of sensor is that it requires a clear area above the liquid nitrogen and it is sensitive to vibration. Another arrangement depends on the condensation of a vapour in basically a gas thermometer. For example a pipe filled to several atmospheres of argon at

room temperature when immersed in liquid nitrogen will result in the pressure being reduced to a vacuum. Such a system can be arranged either to open and close the valve by a piston mechanically connected to the valve seat or to activate an electromechanical switch. A variation on this has been described by Chhatwal[17]. In his arrangement a mercury column was moved along a glass tube by the change in pressure in a gas-filled capsule connected to it. The capsule acted as the level sensor. The column made contact with leads sealed through the glass to activate the valve. These systems can be used to maintain the liquid at a constant level.

However, often it is useful to fill up the container to a much higher level than necessary so that the refilling cycles are not too frequent. One sensor which allows this is a long platinum resistor wound on a vertical former. The resistance value will depend on the length of the resistor covered by the liquid nitrogen. By adjusting the circuit, a threshold resistor value can be chosen to operate the opening of the transfer valve and another to close it and these can be set at convenient liquid nitrogen levels.

A common sensor for the liquid nitrogen level is the thermistor. It is convenient in size and can have a high temperature coefficient of resistance. Two thermistors are normally employed, one mounted above the other. When the minimum level is reached, the lower thermistor changes its resistance and operates the opening of the transfer valve which closes again when the liquid level reaches the upper thermistor. It is important to choose a thermistor with a high temperature coefficient of resistance at 77 K to give reliable operation of the appropriate drive circuit. Normally negative temperature coefficient thermistors are used but Wittstock[18] points out that there are now positive temperature coefficient thermistors available exhibiting a steep characteristic around 77 K which can be used without amplifiers and give rise to smaller control delays. The main problem with thermistors is obtaining sufficient signal when the transistor becomes immersed or emerges from the liquid. Some Dewers, especially cold traps, are very enclosed with only a small filling tube and the gas temperature above the liquid is close to that of the liquid itself. This, coupled with the inherent thermal time constant of the thermistor, can cause delays of minutes with the consequent danger of the level dropping below the critical value. A technique has been employed by Herbert[19] whereby a small heater is wound around each thermistor to supply a small amount of energy all the time. The amount of energy supplied ensures that as soon as the thermistor bead is wholly above the liquid nitrogen it heats up. Wire resistance sensors with a heating current flowing through them could be similarly adapted. Such systems give control delays of less than a few seconds.

The transfer valve must be suitable for low-temperature liquid operation and the transfer tube must be heat insulated. Often this is provided by enclosing the tube in a vacuum jacket. Even so, heating of the liquid initially during transfer can cause the liquid to partially vaporize, until the tube cools down. To minimize the disturbances within the liquid nitrogen container from this gas–liquid mixture, a phase separator can be incorporated close to the container. Basically this is an enlargement of the pipe having a gas outlet hole as illustrated in *Figure 6.19*. To prevent pressure build up in the transfer tube the transfer valve should be mounted close to the reservoir.

Some systems are available commercially which are semi-automatic in that

they deliver a fixed quantity of liquid to the container at regular intervals. The time interval or the quantity is variable and the operator sets the rate to his requirements basically by trial and error. It is also possible to purchase transfer systems which can be operated manually by a switch and some of the automatic systems provide a manual override.

6.7 References

1. PIRANI, M. and YARWOOD, J., *Principles of Vacuum Engineering*, Chapman and Hall, (1961)
2. HAIT, P. W., *Vacuum*, **17**, 547, (1967)
3. EDWARDS, T. J., BUDGE, J. R. and HAUPTLI, W., *J. Vac. Sci. Technol.*, **14**, 740, (1977)
4. DE CHERNATONY, L., *Vacuum*, **27**, 605, (1977)
5. BUCHTER, H. H., *Industrial Sealing Technology*, J. Wiley & Sons, (1979)
6. ROTH, A., *J. Vac. Sci. Technol.*, **A1**, 211. (1983)
7. WHEELER, W. R. and CARLSON, M., *Trans. 8th Vac. Symp. and 2nd Internat. Congress 1961*, Pergamon Press, 1309, (1962)
8. HEAD, P. V., MARTIN, D. M., ALLISON, W. and WILLIS, R. F., *Vacuum*, **32**, 639, (1982)
9. SAKAI, I., ISHIMARU, H. and HORIKOSHI, G., *Vacuum*, **32**, 33, (1982)
10. COENRAADS, C. N. and LAVELLE, J. E., *Rev. Sci. Instrum.*, **33**, 879, (1962)
11. BUDGEN, L. J., *Vacuum*, **32**, 627, (1982)
12. RAJ, K. and REISER, C., *Laser Focus*, 15th April, 56, (1979)
13. ALPERT, D., *J. Appl. Phys.*, **24**, 860, (1953)
14. BÄCHLER, W. and WIKBERG, T., *Vacuum*, **21**, 457, (1971)
15. WHEELER, W. R., *J. Vac. Sci. Technol.*, **13**, 503, (1976)
16. LEEFE, S. and LIEBSON, M., *Rev. Sci. Instrum.*, **31**, 1353, (1960)
17. CHHATWAL, H. L., *J. Physics E., J. Sci. Instrum.*, **5**, 541, (1972)
18. WITTSTOCK, J., *Vacuum*, **29**, 451, (1979)
19. HERBERT, B. J., *Brit. Patent* 1,137,704, (1968)

Chapter 7

Systems and applications

7.1 Introduction

There are a wide variety of applications requiring ultrahigh vacuum systems. The three major areas are particle accelerators, space environment studies and surface studies, but there are several other applications where freedom from gas monolayers or hydrocarbon contamination are essential. Most of the applications are in research and development areas but there are now some production areas, especially in the electronics industry, where pressures below 10^{-6} Pa are required. Examples of such applications are the laying down of epitaxial layers by molecular beam techniques for special semiconductors and the formation of caesiated gallium arsenide photocathodes (see Section 7.5).

Because of the diversity of requirements most systems have to be custom built, often by the user. Particle accelerators and space environmental chambers are extremely expensive systems and the user can afford to dictate the specification and the design of the components. For surface science studies several techniques have become established, such as LEED (low energy electron diffraction) and SIMS (secondary ion mass spectrometry) and specialized vacuum engineering companies design and manufacture the complete equipment. For many of the applications, however, the user must design and build the system using off-the-shelf components chosen from the range available from the vacuum equipment manufacturers. It is to these latter users that this chapter on systems is mainly addressed.

7.2 Systems requirements

Before embarking on building a vacuum system the vacuum engineer should be very sure of his requirements. Perhaps one of the major mistakes made in setting up an ultrahigh vacuum system is over designing. The engineer incorporates much larger pumps than are necessary or too many valves and ports, with a result that the system is more complex and costly than is warranted without any improvement in performance. Indeed in some cases the performance may be degraded by the extra complexity.

In defining the specification, one first needs to know what sort of vacuum

chamber is required, its size, the number and nature of feedthroughs to be incorporated and its orientation. It is important to have a reasonable assessment of the outgassing rate of the chamber particularly when the experiment or process is operating. Having decided on the working chamber, the vacuum environment must be specified, since this defines the pumping requirements. For example, what ultimate pressure is required and does it matter what the residual gas is? Will there be a need to raise the pressure during processing above 10^{-6} Pa, for example for sputter cleaning or chemical reactions and if so what gas will be used? Is it essential to exclude even the minutest amount of hydrocarbons? What pump down time is required? The vacuum measurements that will be necessary should also be considered. Will total pressure measurements suffice or will a gas analyser be essential? What pressure range should be covered and how accurate do the measurements have to be?

Also one should consider the 'plumbing' requirements imposed by the positioning of the chamber relative to pumps and gauges, the need to isolate sections of the system with valves, the provision of gas inlet systems and the demountable connections for servicing or replacement of components. There may be special requirements such as freedom from magnetic fields, freedom from vibration, low-temperature operation or high-voltage electrical insulation. Finally, there are some general considerations which apply to all ultrahigh vacuum systems, such as the materials to be used in its fabrication and the need to provide appropriate baking facilities.

To complete the system one will probably need to build in safety facilities to guard against eventualities such as mains electrical failure or malfunctioning of components. In all it adds up to a formidable list of design considerations which may send all but the stout hearted to one of the number of small companies that have been set up to design and custom build vacuum systems using bought-in components.

One thing is certain, however, to build the system without a clear idea of the requirements is to court disaster and can prove a costly exercise. As an example, in the author's own laboratory an expensive ultrahigh vacuum bell-jar system was purchased in the 1960s for some surface experiments requiring a background pressure of 10^{-8} Pa, the figure guaranteed by the manufacturer. However, when the experimental apparatus involving mechanical movements and evaporation sources was installed in the bell jar, the extra gas load made it impossible to achieve a pressure much lower than 10^{-6} Pa. In the author's experience this was not an isolated case and considerable effort and expense is often incurred because not enough thought had been given to the original requirements.

Having defined the requirements, the system design to achieve the specification can be considered and it is here that some guide lines on the design procedure can be helpful. Unfortunately the required design is unlikely to be a simple matter of adapting a standard commercial system. The number of criteria involved would make it impossible to provide even basic systems which could be adapted for all applications and for any one application there may be several designs which will be satisfactory. To devote this chapter, therefore, to a catalogue of ultrahigh vacuum systems, either available or described in the literature, would serve little purpose. However, it is hoped that the following discussion on the basic design procedures, together with some

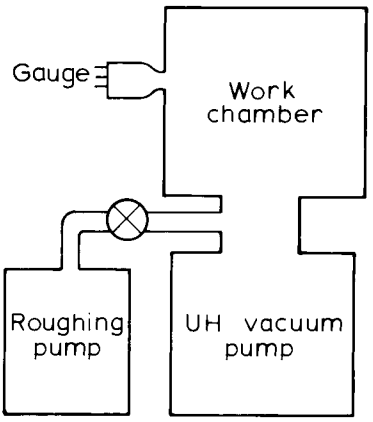


Figure 7.1 Basic pumping system

practical examples will at least help the vacuum engineer to avoid some of the pitfalls which seem to bedevil ultrahigh vacuum system design.

7.3 Basic pumping considerations

Consider the basic pumping system, shown schematically in *Figure 7.1*, consisting of a vacuum chamber in which experiments are to be carried out, the ultrahigh vacuum and roughing pumps, appropriate gauges and isolation valves. At this stage the types of component are not considered.

The evacuation of the chamber to the ultimate pressure from atmospheric pressure can be considered to occur in two distinct steps. The first step involves the removal of the bulk of the volume gas by the roughing pump, which takes the pressure down to say 10 Pa. The second step involves not only removing the rest of the volume gas but also the gas adsorbed on the surfaces, or absorbed within the material, of the chamber and other vacuum components. At this stage the ultrahigh vacuum pump would be brought into operation. In the first step, the influx of gas from the surfaces, etc. is negligible compared with the volume of the gas in the system and the pumping equation (Equation (1.34)) reduces to

$$P_t = P_A \exp(-St/V) \tag{7.1}$$

or

$$t = \frac{V}{S} \ln(P_A/P_t)$$

where P_t is the pressure at time t, P_A is atmospheric pressure (i.e. P_t at $t = 0$), S is the pump speed at the chamber and V is the chamber volume.

In *Figure 7.2*, P_t has been plotted against t on a log-linear scale for unit volume of 1 m^3 pumped at unit speed of 1 $m^3 s^{-1}$. P_A is taken as 1.013×10^5 Pa. The pump-down time to reach a pressure P_t for any system is obtained by multiplying the corresponding time by V/S, assuming that the pump speed remains constant over the pressure range being considered. It should be noted, however, that S is the equivalent pump speed at the chamber and that if the

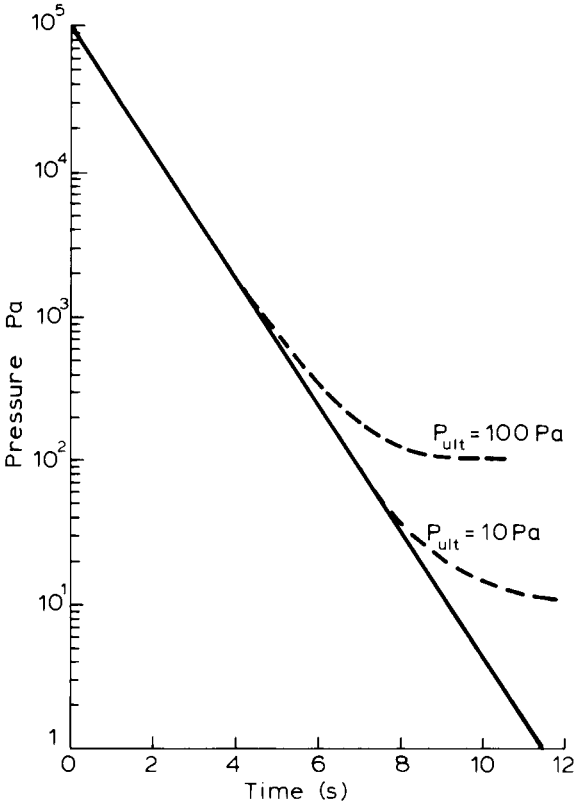


Figure 7.2 Pump-down curve for 1 m^3 volume with a pump having 1 $m^3 s^{-1}$ pump speed. Dotted curves show the effect of gas influx of various levels

pump is connected to the chamber with a pipe of conductance C then t must be multiplied by

$$\left(\frac{1}{S_p}+\frac{1}{C}\right)V$$

where S_p is the actual speed of the pump. At atmospheric pressure the gas flows under viscous flow conditions and the conductance is dependent on the pressure in the tube. For a long tube it is given by

$$C=\frac{\pi}{8}\frac{r^4}{L\eta}\bar{P} \tag{7.2}$$

where r is the pipe radius, L its length, η the viscosity of the gas and $\bar{P}$ the average pressure in the pipe. As the pressure falls the conductance reduces until molecular flow conditions are reached when the conductance for a long tube is given by Equation (1.24)

$$C=\frac{4}{3}\frac{r^3}{L}\left(\frac{2\pi kT}{m}\right)^{1/2} \tag{7.3}$$

Consider now a practical case of a bell jar system pumped down to 10 Pa with a rotary pump. Taking the bell jar as a cylinder of 30 cm diameter and 40 cms high, the volume to be pumped will be 0.028 m^3. It is assumed that a 2.5 $m^3\,h^{-1}$ (7×10^{-4} $m^3 s^{-1}$) pump is used connected to the bell jar by a pipe of 15 mm diameter, 0.5 m long. The initial conductance under viscous flow conditions can be deduced by assuming an average pressure of 0.5×10^5 Pa and using η for nitrogen at 20°C as 1.80×10^{-5} Pa s^{-1} (1.80×10^{-4} poise). This gives the conductance of the pipe as 7.0 $m^3 s^{-1}$ which is large compared with the pump speed. It will reduce linearly with pressure as the gas is pumped out but will only be comparable with the pump speed when the pressure has dropped to 10 Pa where it will have fallen to the molecular flow value of 8×10^{-4} $m^3 s^{-1}$. Thus, for most of the pumping cycle the multiplication factor will be $V/S_p = 40.3$ and only rise to 75 when 10 Pa pressure is reached. From *Figure 7.2*, $t = 9.2$ s so that for our practical system we might expect a pump down time of 6 to 8 min. To pump down a further order to 1 Pa might take nearer a half-hour, since, apart from the lower conductance, the pump speed at pressures below 10 Pa drops fairly markedly.

At 10 Pa the ultrahigh vacuum pump, which usually has a much higher pump speed, is brought into operation and the connection between the roughing pump and the bell jar is valved off. The influx of gas Q_I as a result of desorption from the chamber walls, etc. now becomes significant and the full pumping equation must be taken *viz.*

$$P_t = \frac{Q_I}{S} - \left(\frac{Q_I}{S} - P_A\right) \exp\left(-\frac{S}{V} t\right) \tag{7.4}$$

If Q_I remains constant then the ultimate pressure is given by

$$\frac{Q_I}{S} = P_{ult} \tag{7.5}$$

and we can rewrite Equation (7.4) as

$$\frac{P_t - P_{ult}}{P_A - P_{ult}} = \exp\left(-\frac{S}{V} t\right) \tag{7.6}$$

The effect of this is shown in *Figure 7.2* for the unit volume case by plotting the curves for different values of Q_I. Pump-down time to a value close to the ultimate pressure is still fast but the pressure reached is dominated by Q_I and the important criteria is whether Q_I can be reduced and if so at what rate.

Returning to the example system, it is assumed that an ultrahigh vacuum pump is employed with an effective speed of 0.25 $m^3 s^{-1}$ and that the only gas influx comes from the chamber walls. For an unbaked system the degassing rate is likely to be around 10^{-4} Pa $m^3 s^{-1} m^{-2}$ so with the chamber surface area of 0.52 m^2, Q_I will be 0.52×10^{-4} Pa $m^3 s^{-1}$ and the ultimate pressure will only be 2×10^{-4} Pa. For a thoroughly degassed system the desorption rate decreases to 10^{-10} Pa $m^3 s^{-1} m^{-2}$ and the ultimate pressure reached will be 2×10^{-10} Pa. The pump-down time to reach 2×10^{-10} Pa will therefore depend on the speed with which the degassing rate can be reduced. Thus it can be stated that the pump-down time taken to attain ultrahigh vacuum is dominated by the time required to reduce the outgassing rate of the vacuum chamber and other components.

From Chapter 2 we saw that to a large extent the degassing rate depended on the material and the cleaning and stoving processes applied before being assembled in the vacuum system. Nevertheless, with an unbaked vacuum system the reduction in the degassing rate to that required for ultrahigh vacuum can still take several hundred hours, even with nominally clean stoved material.

Initially the main source of gas is that adsorbed on the surfaces. This was discussed in Section 1.5.4 where the desorption time, τ, was shown to be exponentially dependent on the reciprocal of the temperature.

$$\tau = \tau_0 \exp(E/RT) \tag{7.7}$$

where τ_0 is a constant and E is the binding energy between the molecules of the gas and the solid.

For physical adsorption, E is around 12 kJ mol^{-1} and τ varies from 160 τ_0 at room temperature to 8 τ_0 at 450°C. For chemical adsorption, E may be much higher, say 210 kJ mol^{-1}, and τ will then vary from 10^{36} τ_0 at room temperature to 10^{15} τ_0 at 450°C. Theoretically it has been suggested that τ_0 should be of the order of 10^{-13} s but practical values recorded have been much higher. Whatever the value of τ_0, it is clear that baking to 450°C has a pronounced effect. Provided the gas is physically adsorbed, the removal of surface gas should be fairly rapid, i.e. minutes rather than hours (cf. *Figure 1.4*). For chemisorption, for example as an oxide layer, the desorption time will be considerably longer and baking to 450°C becomes essential.

Having removed the surface gas the problem is then to remove the gas absorbed within the bulk of the material making up the vacuum system. Again this was discussed in Chapter 1. The process is controlled by diffusion and from Equation (1.56) the degassing rate per unit volume of an infinite slab, through unit area exposed to the vacuum was given by

$$Q_B = c_0 \left(\frac{D}{\pi t}\right)^{1/2} \tag{7.8}$$

where c_0 is the initial concentration of gas in the material and D is the diffusion rate, which is temperature dependent. To get some idea of the quantities involved, D for hydrogen in iron is of the order of 10^{-13} m^{2}s^{-1} at room temperature, rising to 10^{-8} m^{2}s^{-1} at 450°C. The respective values for nitrogen in iron are 10^{-35} m^{2}s^{-1} and 10^{-16} m^{2}s^{-1}. The quantity of gas trapped within untreated metal can be as high as 100% of the volume at STP or 10^5 Pa m^3 m^{-3}. Equation (7.8) applies to an infinitely thick slab of material. For a slab of finite thickness with one surface at atmospheric pressure and the other at vacuum pressure, the equations become more complex. However, as an approximation, one can consider the gas as coming from a depth, d, of an infinite slab and Equation (7.8) can be written as

$$Q_B = \frac{c_0}{d}\left(\frac{D}{\pi t}\right)^{1/2} \tag{7.9}$$

If we assume in our example system that the wall thickness is 2.5 mm then, using the values of D and c_0 given above, the rate of degassing for nitrogen even in the first few seconds would be less than 10^{-10} Pa m^{3}s$^{-1/2}$m^{-2} and should not be a problem. On the other hand the value for hydrogen would be

11 orders higher at 10 Pa $m^3 s^{-1/2} m^{-2}$. In practice, the gas absorbed within the metal is there as a result of the smelting processes and the concentration of hydrogen will be much lower than, for example, oxygen. Nevertheless, values between the extremes for nitrogen and hydrogen are likely to occur and the evidence is that, after the initial outgassing, the rate settles to a value which is inversely proportional to the square root of the pumping time. This is a clear indication that diffusion of the bulk gas becomes the controlling mechanism.

Although the theoretical Equation (7.9) does not give a quantitative prediction of the degassing behaviour of a vacuum system, it is in agreement with general experience. For example, to reduce the ultimate pressure by, say, three orders could be expected to take 10^6 s, i.e. two or three hundred hours. However, this can be considerably reduced by increasing D during the pumping period. Raising the temperature to 450°C increases D by several orders of magnitude and the reduction in ultimate pressure may be attained within a few hours.

In summing up it can be said that pumping down to 10^{-4} Pa can be achieved fairly rapidly, but to get from 10^{-4} Pa to 10^{-6} Pa will take considerably longer, hours rather than minutes. The process can be speeded up by baking to 450°C, but even then with vacuum stoved components it can still be a matter of an hour or so. If an ultimate pressure of 10^{-8} Pa is required then a prolonged bakeout will be needed of at least 8 hours. If baking at 450°C is not possible then it could take several weeks to achieve the ultimate pressure.

The above premises are based on the assumption that the only influx of gas is due to desorption of gas from the vacuum components. If there is a leak, or gas permeation through the walls, or a material with a high vapour pressure is included in the system, then no amount of pumping or baking will reduce the influx. It is essential, therefore, to ensure that gas influx from these latter causes are below that necessary to achieve the specified ultimate pressure. In Chapter 2 the permeation of gas through various materials was discussed and in Chapter 8 the methods of detecting small leaks are described. A cause of gas influx which is often present and difficult to detect are faults in the material of the vacuum envelope which render it somewhat porous. The leak, through individual faults, may be below the sensitivity of a leak detector but summed over the surface it may amount to an unacceptable influx. Also one must guard against virtual leaks such as a screw in a blind hole, where gas may be trapped and slowly leak out when under vacuum.

Before leaving this section, a further point should be made and that is the relationship between the ultrahigh vacuum pump speed and that of the backing pump, for a diffusion or turbomolecular pumped system. The quantity of gas pumped per second by the ultrahigh vacuum pump is given by

$$Q_u = S_u P_u \tag{7.10}$$

where S_u is the effective pump speed and P_u the pressure in the system. Similarly the gas quantity pumped by the backing pump will be

$$Q_B = S_B P_B \tag{7.11}$$

To ensure that the backing pump can cope with the throughput of the high vacuum pump, Q_B must be equal to or greater than Q_u and thus

$$S_B \geqslant \frac{S_u P_u}{P_B} \tag{7.12}$$

Under equilibrium conditions, when the ultimate pressure of, say, 10^{-8} Pa has been reached, we see that for a backing pressure of 10 Pa the backing pump need only have a pump speed of $10^{-9}\,S_u$. However, when the ultrahigh vacuum pump is first switched on, the pressure in the system is equal to the backing pressure and, with a lower pump rate for the backing pump, the backing pressure will start to build up exponentially towards $P_{BO}(S_u/S_B)$ where P_{BO} is the backing pressure value at which the ultrahigh vacuum pump was brought into operation. The exponential time constant for this build up will be V_B/S_B where V_B is the volume between the two pumps. Since there is an upper limit for the backing pressure, above which the diffusion or turbomolecular pump will not operate, normally of the order of 10^3 Pa, the backing pump pressure must not be allowed to rise to this critical level. Fortunately, as the pressure rises in the backing line, so too is the pressure in the ultrahigh vacuuum system reduced by the pumping. This means that Q_u will be diminishing exponentially with a time constant of V_u/S_u where V_u is the volume of the ultrahigh vacuum chamber above the pump. If this exponential time constant is similar to V_B/S_B, then the backing pressure will not rise appreciably and this can be taken as an upper value for the backing pump speed, i.e.

$$S_B = \frac{V_B}{V_u} S_u \tag{7.13}$$

Although V_B should be much less than V_u, clearly it is unlikely to be smaller by a factor of 10^9, so that S_B will need to be considerably higher than that required under equilibrium conditions. As a rule of thumb, a value of $S_B = 10^{-2}S_u$ is considered a satisfactory and completely safe figure for most applications but the relative values of the volumes V_B and V_u should be carefully borne in mind when designing the system. Normally, manufacturers will specify the minimum backing pump speed for their diffusion or turbomolecular pumps or recommend the rotary pump that should be used.

7.4 Approaches to system design

Having considered the basic factors controlling the evacuation of a system, we can now consider how these factors affect practical designs. It is not intended here to give design data for the construction of specific systems but rather to outline some of the principles involved in deriving such design data. It is convenient to discuss these principles under headings defined by the type of pump or pump combinations which are to be used, although some of the principles discussed may be equally applicable to several or all types of pumping systems.

7.4.1 Diffusion/rotary pumped systems

A schematic diagram of a typical ultrahigh vacuum system using a diffusion pump backed by a rotary pump is illustrated in *Figure 7.3*. The dotted area represents that part of the vacuum system which can be baked and the vacuum valves V_1, V_2 and V_3 should therefore be bakable. V_2 and V_3 can be all-metal valves. The high vacuum valve, V_1, has to have a conductance which is greater

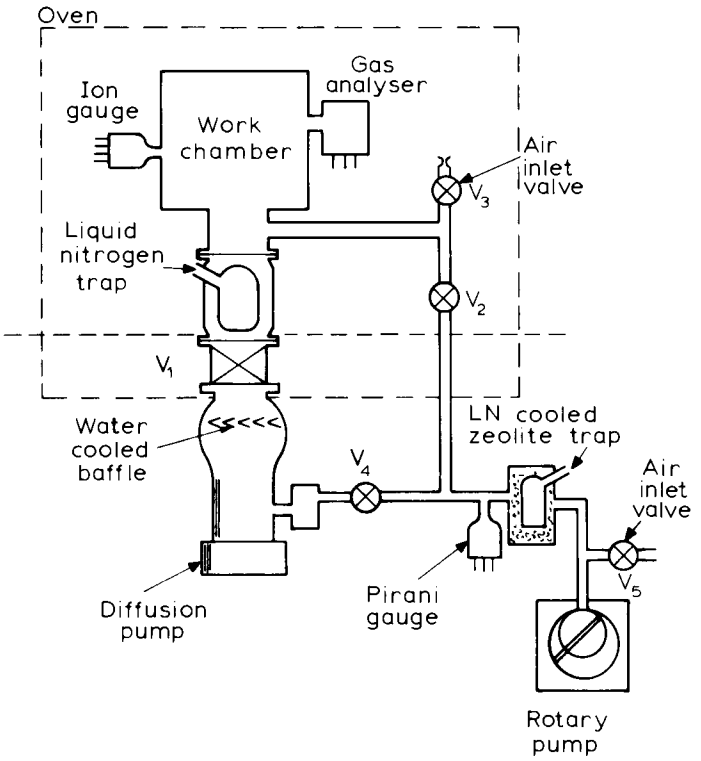


Figure 7.3 Typical diffusion pump/rotary pump system

than the pump speed of the diffusion pump and for such a large conductance, the valve in an all-metal construction would be very expensive. Therefore, for most applications, a Viton A sealed gate valve or plate valve will be used. This can be baked up to 250°C in the open position. If a higher temperature bake is required then the oven would have to be raised above the valve V_1, to the position shown by the dashed line. For some applications V_1 can be omitted but in this case the system should not be let up to air until the pump has cooled down. With the new pump fluids, such as polyphenyl ether and naphthalene based oil, it is claimed that back-streaming can be eliminated by a well designed water-cooled baffle. For ultrahigh vacuum a liquid nitrogen cooled trap is often incorporated as an extra precuation. It also serves as a cryopump for any water vapour or carbon dioxide which may be evolved from the system. Alternatively, in some pumps a cryopanel can be incorporated which sits above the water-cooled baffle and is cooled via a copper rod connected to a liquid nitrogen reservoir[1]. A zeolite trap to prevent backing pump oil getting into the system is essential.

The procedure for pump down should be along the lines of the following description. First, the system is roughed out via the valve V_2, with all other valves closed until a pressure of around 10 Pa is reached. This ensures that large quantities of gas are not pumped through the diffusion pump, with the possibility of oil being swept out with the gas. V_4 is then opened and the diffusion pump switched on. When it is operating, V_2 is closed and V_1 is

opened. When the pressure has dropped to around 10^{-5} Pa, the oven is switched on and the system, including the liquid nitrogen trap, is baked to the highest temperature allowed by the valves. Several hours baking will be required. The oven is then allowed to cool and the liquid nitrogen trap filled. A further bake of that part of the system above the liquid nitrogen trap may improve the vacuum, especially if a higher temperature can be employed.

Essential to the diffusion pump–rotary pump system is protection against failure of the electrical mains and of the supply of water for cooling. The former is normally taken care of by making valves V_1, V_4 and V_5 solenoid operated, so that V_1 and V_4 close and V_5 opens when the electricity is cut off. The whole system—pumps, valves and also the ovens if they are on—must be controlled by a trip system which does not restart automatically when the electricity is reconnected. The failure of the water supply, or air cooling if it is used, is taken care of by a thermal switch on the pump, which then switches off the pump heater and closes valve V_1 when the pump temperature rises above a certain level.

Common to all systems, the filaments of the ionization gauges and gas analysers need to be protected against the ingress of air and most gauge and analyser supplies incorporate a cut-out, which switches off the filament when the pressure rises above a set value. If the ingress of air is large, then V_1 should also be closed to prevent a rush of air through the hot diffusion pump. In *Figure 7.3* a Pirani gauge is shown in the roughing line. This not only indicates the backing pressure but can also be used as the safety switch to close V_1, when the pressure rises significantly above the normal backing pressure. It is also useful for indicating when the pressure has fallen sufficiently, on the initial pump down, for the ion gauge to be switched on. It is important to outgas the ion gauges and analysers after the baking process to ensure their proper functioning.

Finally, a word about the liquid nitrogen cooled trap. Ingress of air will cause condensation of water vapour and carbon dioxide. If the pressure rises to atmospheric, liquid oxygen can be formed in the trap which could be dangerous if swept into the pump. It is therefore important to empty the liquid nitrogen from the trap should the system be let up to air, either accidentally or as part of the processing procedure. Valve V_1 can be mounted above the liquid nitrogen trap provided the valve and activating mechanism can be fully baked. This would overcome the problem, since the trap could then be isolated should the system pressure rise to atmospheric through any malfunctioning.

Because the diffusion–rotary pumped system can handle a large gas throughput, it is particularly suitable for systems such as sputtering chambers where large gas loads may occur.

7.4.2 Turbomolecular pumped systems

Turbomolecular pumped systems are similar in many ways to diffusion pumped systems and indeed some of the single flow ('vertical') types of pump are designed to be fitted as a direct replacement for the equivalent diffusion pump. The turbomolecular pump is normally backed with a rotary pump and a similar valving system to that of *Figure 7.3* can be employed. The system differs from that using the diffusion pump in that the water-cooled baffle and liquid nitrogen trap are not required. This is because the compression ratio for

the heavy molecules is so high that virtually none can back diffuse. However, as pointed out in Chapter 3, most of the turbomolecular pumps have oil lubricated bearing which are exposed to the system and if the pump is stopped for any reason whilst it is under vacuum, the bearing oil will diffuse into the system. The pump must, therefore, be protected by closing valve V_1 and letting air into the pump in the event of an electrical mains failure. The pump should not be vented on the backing pressure side lest the flow of gas through the pump entrains the lubricating oil molecules. Some pumps are provided with a venting port between the compression stages so that the gas is evenly distributed between the high vacuum and backing pressure sides of the pump. The venting valve does not then need to be bakable.

The main limitation of a turbomolecular pump is the relatively low compression ratio for hydrogen which results in a high hydrogen residual gas component. This hydrogen partial pressure can be reduced by two methods. First, by reducing the hydrogen partial pressure in the backing line and secondly by combining the pump with other ultrahigh vacuum pumping systems. To ensure a low pressure of the hydrogen in the backing line it is important to employ an adequate rotary pump (see Section 7.3). There is also evidence that the rotary pump oil can affect the results[2]. Alternatively the turbomolecular pump can be backed by a diffusion pump with its associated backing pump. Perhaps the best solution, however, is to combine the turbomolecular pump with a titanium sublimation pump connected on the high vacuum side. The sublimation pump has a high speed for pumping active gases, particularly hydrogen, and is relatively cheap. It would not need to be continuously run.

Most turbomolecular pumps can be baked to a temperature around 100°C, which can be achieved with heating tapes or the pump may be provided with a suitable heater (see *Figure 3.7*). The pump down and baking procedure is basically similar to that for the diffusion pumped system.

7.4.3 Ion pumped system

Because the ion pump holds the pumped gas within the pump body it does not require to be backed by a continuously running backing pump. It will not operate, however, until the pressure is reduced to at least 10 Pa, so a method of reducing the pressure from atmospheric is still required. Since the main attribute of the ion pump is complete freedom from hydrocarbons, using an oil-sealed rotary pump for this purpose rather defeats the object. Although there are now available oil-free rotary pumps and diaphragm pumps for clean vacuum application, they do not pump down to a low enough pressure for ion pumping. Normally, therefore, the ion pump is used in conjunction with sorption pumps, a typical arrangement being illustrated in *Figure 7.4*. With the ion pump there is no need for a cold trap or a baffle, nor does any precaution have to be taken against electrical failure although if the pressure rises restart might overheat the pump. In general there is no need for an isolation valve between the pump and the system. The absence of the cold trap and isolation valve means that full use can be made of the available pump speed to the extent of mounting the pump within the working chamber in some designs.

When commencing to use the system, the first task is to activate the sorption pumps by heating them to 250°C whilst open to the atmosphere via valve V_1 or

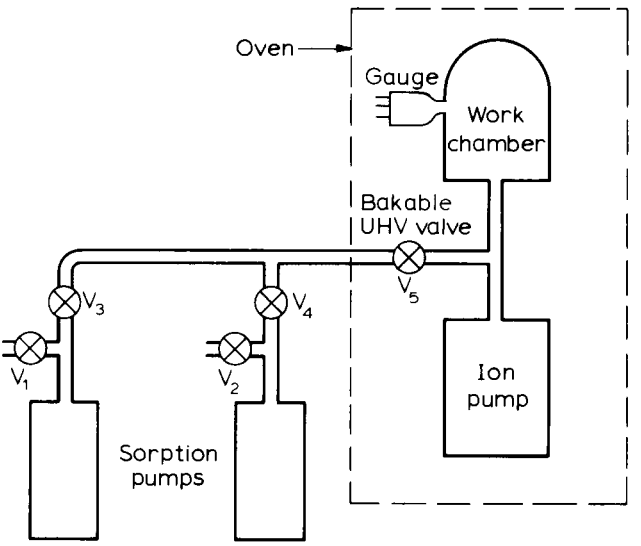


Figure 7.4 Simple ion pumped system

V_2, the valves V_3 and V_4 to the system being closed. This process drives off the water vapour which will have accumulated in the pumps when exposed to air for any length of time. After this activation process V_1 and V_2 are closed and the pumps refrigerated with liquid nitrogen. The gas already in the pumps acts as a conducting medium allowing the zeolite pellets, which are notoriously bad thermal conductors, to be efficiently cooled. The first sorption pump is then opened up to the system via V_3 and the pressure will fall to around 1 Pa.

At this stage V_5 is also open and the system can be gently baked to 120°C to speed up the evolution of water vapour from the surfaces. The first sorption pump is then isolated from the system by closing V_3 and the second 'unsaturated' sorption pump connected to the system by opening V_4. The system can now be baked to 250°C with the ion pump switched on. This baking not only helps to outgas the system but also ensures that the ion pump is cleaned and will operate at its maximum pump speed. This is known as a regenerative bake. It is considered that the high temperature increases the sputter yield and releases trapped gas from the surfaces. The sorption pumps can be valved off during this bake. The pressure should drop after this bake to at least 10^{-3} Pa. V_5 is then closed and further baking of the system to 450°C, not the ion pump, which should not go above 250°C, should result in ultimate pressures $< 10^{-8}$ Pa. To take care of any subsequent active gas evolution, a titanium sublimation pump can be added to the system to act as a booster pump. Of course it will not pump any of the inert gases.

Both the sorption pump and the ion pump have a limited gas-load capacity, especially for inert gases, and this can be a problem for systems where high loads of inert gas are probable. The capacity of the sorption pumps can be enhanced by a pre-exhaust using an auxiliary pump such as an oil-free mechanical pump. The sorption pumps are activated by baking at this reduced pressure of typically $< 10^{4}$ Pa and as a consequence the neon residual pressure is lower as well as the capacity for active gases being increased. Flushing the

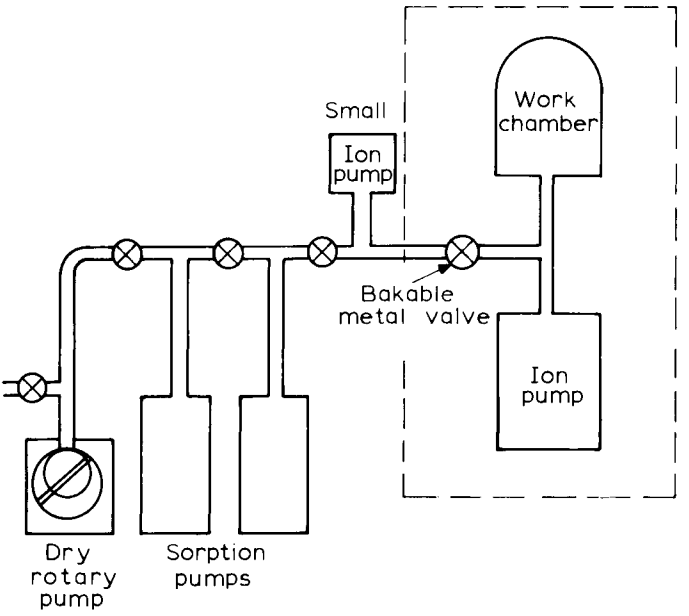


Figure 7.5 An improved ion-pumped system with auxiliary pumps

system through with dry nitrogen before sorption pumping will also improve the ultimate pressure reached by the sorption pumps. The capacity of the ion pump can be enhanced by using an auxiliary ion pump, which need not be large, to pump the gas during the regenerative bake. This leaves the main ion pump to evacuate the system to the ultimate pressure. Such a system is illustrated in *Figure 7.5*

An alternative solution for handling high gas loads is to use a hybrid system of more than one pump, especially to combine an ion pump with a turbomolecular pump. The system illustrated in *Figure 7.6* is useful for long-term experiments, where there is the possibility of large gas loads either by accident or by design. The gas handling capacity of the turbomolecular pump takes care of the gas loads and gives a low pressure for the 'start-up' of the ion pump which ensures the attainment of a low ultimate pressure. A large conductance bakable valve, V_1, is required between the two pumps as well as a metal backing pump valve.

7.4.4 Cryopumped systems

Cryopumps are akin to ion pumps in terms of system design. Since there are no contamination problems, advantage can be taken of their high pump speed and, in a similar way to ion pumps, the cryopanels can be installed in the actual work chamber. For large systems such as space simulation chambers, where the pumps are custom designed, this principle of integral pumps is exploited. However, for smaller systems the vacuum engineer relies on commercial cryopumps based on refrigerating machines which are connected on to the system as an appendage.

Because of their limited capacity, the system must be pre-evacuated by a

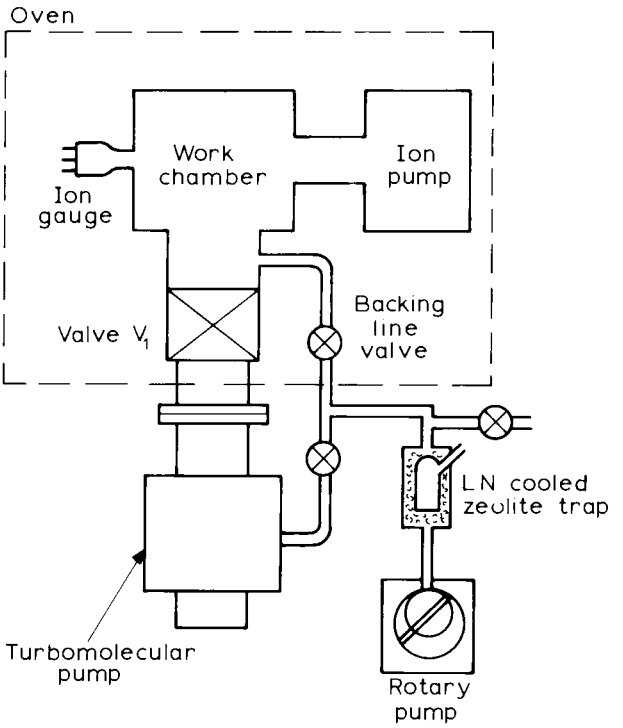


Figure 7.6 A hybrid system for handling high gas loads

backing pump. For contamination free pumping, the sorption pump is an obvious choice for pre-evacuation and for many systems it is satisfactory. However, if a large amount of inert gas such as argon is to be pumped this combination could be less satisfactory. Although the cryopanel will readily pump the gas, when the cryopump needs to be regenerated, there will be difficulty in pumping the gas away with the sorption pumps. The alternative of using a well trapped rotary pump is often advocated. Because the backing pump is only required for the initial pump-down period, contamination is minimal. It can be reduced further by operating the backing pump in the viscous flow region, i.e. above a pressure of 200 Pa. However, apart from the extra gas load, starting the cooling down process of the cryopump at 200 Pa can result in condensable gases such as CO_2 and water vapour being trapped on the cryopanel rather than the surrounding shield long before the final temperatures have been reached. Such an eventuality will effectively reduce the amount of charcoal coated on the cryopanel available for sorption pumping hydrogen and helium at a later stage. The problem can be alleviated by introducing a valve between the pump and the system. *Figure 7.7* illustrates a suitable system. It should be pointed out, however, that if the high vacuum valve is to be bakable it could be expensive if it is not to impede the pump speed.

Because of their high pump speed, cryopumps are often used with systems which cannot be easily baked to a high temperature. To reduce outgassing from the surfaces, the work chamber walls may be cooled by liquid nitrogen. If

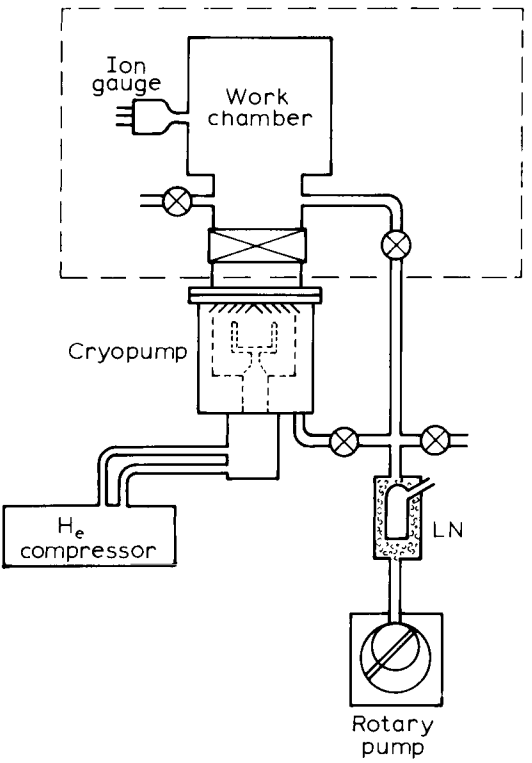


Figure 7.7 A cryopumped system using a rotary pump for roughing out

this is the case, then care must be taken in interpreting ion gauge readings. Even if the walls are not cooled, siting of the gauge in a cryopump system is important, particularly whether or not it has a direct view of the pump. The ion gauge measures density not pressure and if the gas in the gauge is at a temperature T_1 whilst that in the rest of the system is at T_2, then the ion gauge pressure reading $P_i = P(T_2/T_1)^{1/2}$ where P is the pressure in the system. The various ramifications of this premise for a cryopumped system have been discussed by Haefer[3].

Baking a cryopump system can present difficulties. In most pumps the radiation heat load is limited and baking the system at 450°C, for example, can result in the temperature in the pump rising above its operational range. A water-cooled baffle can be inserted between the pump and the system as a radiation shield but this seriously reduces the pump speed. An alternative has been suggested by Kubiak *et al.*[4] whereby a baffle is swung into position during the baking period and moved out of the way thereafter. Their design is similar to a quarter-swing valve. In position it reduces the pump speed by about 75%.

Since the pump speed for hydrogen is relatively low, incorporating a titanium sublimation pump can be an advantage. The cryopump can also be used in conjunction with other ultrahigh vacuum pumps such as an ion pump or a turbomolecular pump. It should be appreciated, however, that the

cryopump is expensive and its inclusion in a hybrid system will only be justified if its attributes are essential.

7.4.5 Automatic control

Whatever pumping method is employed, obtaining the optimum vacuum environment from the equipment depends on carrying out a suitable processing sequence of pumping and baking. The steps in the sequence involve bringing pumps in and out of operation at specific times, manipulating valves and switching the ovens and controlling their temperature. These steps and their timing are dictated by the state of the system at any one time, which is controlled by such factors as pump-down speed, outgassing rate, oven warm-up time, etc. Such a procedure lends itself to automatic control by a microcomputer, which would not only be able to carry out the most suitable sequence but could be programmed to cope with supply failure or malfunctioning of components. It might also be extended to cover any other processing required in the work chamber. Such automatic sequence control systems are most suitable for high vacuum equipment, where the valves can be electromechanically operated at very little extra cost. Indeed many of the large equipments on the market, such as coating units, ion milling systems and leak detectors are offered with automatic control of the pumping and measuring units.

Automatic control of an ultrahigh vacuum system is less easily achieved. Bakable metal valves are generally operated manually. A few manufacturers offer them with hydraulic or electromechanical drive but then the drive mechanism cannot usually be baked. VAT produce pneumatically driven all-metal valves which are fully bakable up to 450°C. However the drive mechanism takes up a lot of space, 120 cm^3 for a 16 mm valve and they are expensive, more than twice the price of a manual operated valve. There would also be difficulties if the pumping sequence involved filling and/or emptying a liquid–nitrogen vessel, for example where sorption pumps or cooled traps are required. The mechanism to carry out such a process could be complex and may require a cryogenerator. There will no doubt be ultrahigh vacuum systems where such components are essential and the expense of automation is justified. Of more interest are the systems working at the top end of the UHV pressure range, around 10^{-6} Pa, where Viton A sealed valves are adequate and liquid nitrogen, if used at all, is kept at a fixed level. The automation of such systems has been greatly simplified in recent years by the availability of programmable sequence control units. Although designed for industrial processing, they are relatively inexpensive and can be adapted for control of a vacuum system. Basically these control units have a number of inputs to which sensors can be connected. Information from the inputs is fed into a microcomputer which is programmed to carry out various sequences, the order and timing of which are controlled by the input data. The output is fed to an array of output terminals as electrical signals which can be used to activate relays or electronic switches.

A typical unit is the ISCOS 20 from Philips, which is specifically designed for machines requiring sequence control. It is a modular system using standard eurocard printed circuitry. The program is written and stored in a non-volatile EPROM (erasable programmable read only memory) which is separate from

the central processing unit (CPU). Because the EPROM has to be programmed via a number of steps it cannot be altered by accident. The CPU can handle up to 256 inputs and outputs and carry out 2045 instructions. The input and output boards, however, carry only 16 terminals each, so a number of cards would be required for full capacity. For most pumping systems a single input and output card would probably be adequate. The input sensors would probably be gauges, spectrometers, thermometers and possibly liquid flow meters. The output could be coupled to pump supplies, gauge controllers, valve drivers and oven controllers.

Typically the program would set the valves in the correct position for start up and switch on the roughing pump. When the pressure reached a certain value, the main pump would be turned on and the valves sequenced accordingly. If a diffusion pump is used, then the presence of cooling water could be first ascertained. Having reached a suitable pressure the oven could be switched on and the system baked. The sequence could also outgas the gauges. The baking time could be controlled by the pressure or be fixed. The oven would then be switched off and, when cool, removed from the apparatus. If the expected pressures were not reached, the system could be programmed to check for leaks by monitoring the pressure against time in the work chamber when sealed off. Having achieved the required vacuum conditions the controller could then be used for carrying out the required processes, for example, evaporation or ion milling in the work chamber. Otherwise it could be set in a monitoring mode, storing a record of procedures and conditions and acting upon any possible malfunctioning. A description of a microprocessor control of a small vacuum system has been given in some detail by Lucas *et al.*[5].

The main advantages of an automatically controlled vacuum system is that it can be left unattended and can be operated by less skilled personnel. Even for the skilled vacuum engineer, it reduces the likelihood of accidents due to human error. It is of particular interest for vacuum equipment used in production or development where repetitive processing is required but even in a research environment it can be an asset in reducing the time spent by the engineer in operating the vacuum system and making measurements.

7.5 Applications

7.5.1 Particle accelerators and space simulation chambers

Particle accelerators present one of the greatest challenges to the vacuum engineer, particularly the storage ring type, where the particles are required to circulate for periods of many hours. To achieve such life times, it is imperative that the particle loss due to interactions with residual gas molecules is minimal and this implies gas pressures in the ultrahigh vacuum region. Achieving ultrahigh vacuum in such rings is no mean task when one considers their size. As an illustration, the electron storage ring for the Synchrotron Radiation Source (SRS) built at the Daresbury Laboratory of the Science and Engineering Research Council (SERC) in the UK, is 30 m in diameter and will be capable, ultimately, of accepting at least 0.5 A of circulating beam current at 2 GeV.

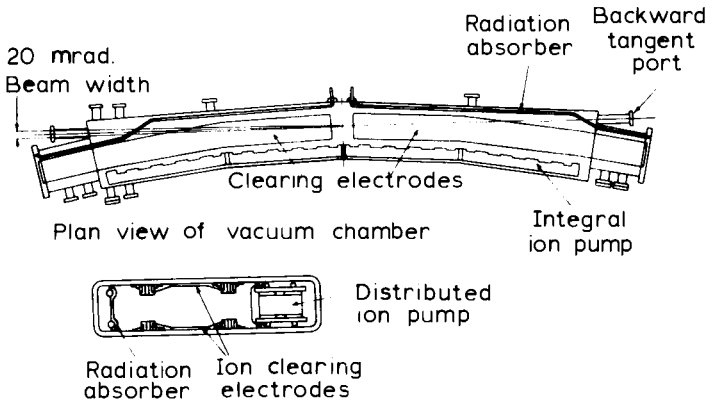


Figure 7.8 Dipole–magnet vacuum chamber for the storage ring of the Daresbury Synchrotron

The storage ring is made up of 16 dipole-magnet vacuum chambers, each about 2.5 m long, interconnected with straight sections which include four accelerating cavities. The main pumping is performed by ion pumps and titanium sublimation pumps (TSPs). There is one 0.4 m^3s^{-1} triode ion pump and TSP connected to each straight section and a distributed ion pump of the differential cathode type along 2 m length of each dipole-magnet chamber. This latter pump uses the fringe magnetic field of the dipole magnets and has an air speed of about 0.3 m^3s^{-1} per chamber. The construction of the chambers is illustrated in *Figure 7.8*.

The ring is roughed out by mobile units each having two sorption pumps with a diaphragm pump to allow vacuum activation of the sorption pumps and roughing to <1 Pa. There are also four all-metal, air-bearing turbomolecular pumps permanently connected to the ring, one for each quadrant, of the type that can exhaust directly to the atmosphere without recourse to a rotary pump. These provide a lower pressure than is available from the sorption pumps alone and facilitate start up of the ion pumps as well as coping with extra gas load in operation and during quadrant bakeout. The main monitoring of the vacuum is carried out using Bayard–Alpert ion gauges and small quadrupole mass spectrometers. The whole system is computer controlled to cover processing and to guard against malfunctions. The ring can be baked out to 200°C with heater tapes and base pressures of better than 5×10^{-8} Pa have been achieved. The system has been described by Trickett[6].

Large though the SRS machine may seem, it almost fades into insignificance when the new storage ring to be built at CERN is considered. This new European machine will accelerate and store electrons and positrons and has been given the title LEP (Large Electron Positron Storage Ring). The machine is scheduled to be commissioned in the late 1980s. The LEP storage ring will be 27 km in length (approximately 8 km diameter) with a cross-sectional form of vacuum pipe roughly 13 cm $\times$ 7 cm. Initially beam energies of the order of 60 GeV are planned but it is hoped to increase this to over 100 GeV by including extra r.f. accelerating cavities or superconducting cavities at a later date.

Machines built for nuclear fusion experiments also involve the evacuation of

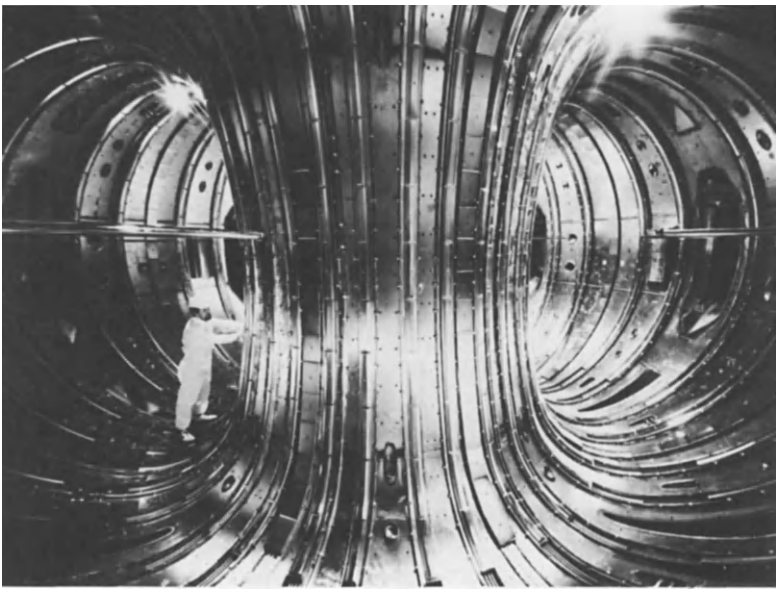


Figure 7.9 Interior of the vacuum vessel of JET Tokamak machine. (Courtesy of JET Joint Undertaking)

large volume vessels to pressures in the ultrahigh vacuum region. *Figure 7.9* shows the interior of the vacuum vessel of JET (Joint European Torus) which was operated for the first time at Culham Laboratory in the UK in June 1983. JET is the largest Tokamak machine in the world and it is hoped that temperatures of the order of 10^{8}°C, hotter than the centre of the sun, will be reached. Initially modest currents of the order of 6×10^{4} A have been passed through a low density of hydrogen gas. Eventually currents around 5×10^{6} A are anticipated with the hydrogen replaced by deuterium and tritium to produce the fusion reaction. Pumping tritium gas, or indeed any radioactive gas, presents special problems because of the health hazard. For those pumps in which the pumped gas is retained within the body of the pump, such as the sublimation pumps designed to work in conjunction with cryopumps in some machines, it is mainly a matter of dealing with the component parts at the end of their operational life. However, for roughing out and dealing with high gas load, for example when regenerating cryopumps, pumps capable of high throughputs are required, i.e. turbomolecular, rotary and mechanical booster pumps. To minimize the release of tritium into the environment, the pumps must be operated as a closed system, in which the exhaust is collected in some form of tritium removal chamber and treated prior to its release into the atmosphere. However, there is another problem with tritium in that it tends to exchange with hydrogen in organic materials, causing radiative contamination and deterioration in the physical properties of the material.

Elastomers and oils can be affected in this way. Much of the tritium can be taken up by incorporating Zr/Al getters in the pumping line but inevitably some contamination of the lubricating oils in the above-mentioned pumps will occur. Special pumps have now been designed for this situation. Turbomolecular pumps have been adapted to stand up to the radiation by a

suitable choice of materials and the oil for the shaft bearings is fed from a sump which can be periodically drained and replenished by remote control[7]. With mechanical booster pumps, there is a danger of gas leakage through the dynamic shaft seal. Budgen[8] has described a design which eliminates the seal by using magnetic coupling (see Section 6.3).

In general, mineral oils, particularly those which polymerize under radiation conditions, are unsatisfactory for pumping radioactive gases. The fully fluorinated fluids are less susceptible to radiation damage and Fomblin* (perfluoropolyether) is recommended as a lubricating fluid for this application. Similar conditions apply to the rotary pumps. It is especially important that all seals and valves used to interconnect with the system and other pumps should be of the all metal variety (see Section 6.2). Also the appropriate lubricating fluid should be used and facilities provided for changing it regularly without danger.

Most particle accelerators and fusion machines allow a certain measure of baking for outgassing the envelope. The surfaces can be cleaned further by passing a glow discharge or the particle beam itself can be enlisted to remove surface gas layers. For space simulation chambers which may be 12 m or more in diameter, baking the chamber is rarely possible. However, to simulate the thermal absorption caused by the emptiness of space, it is necessary to surround the objects under test with a non-reflecting (black) cold surface. Thus virtually all space chambers have surfaces which surround the test object, which are cooled to liquid nitrogen temperatures. Since the main gases desorbed from the stainless steel chamber walls have vapour pressures below 10^{-8} Pa at 77 K, there is little need for a thermal degassing cycle. The obvious choice of ultrahigh vacuum pump to evacuate such chambers is a cryopanel, cooled to liquid helium temperature, incorporated within the space chamber and such designs have been mentioned in Section 3.4.

The consumption of helium in this type of pump can be expensive and for the more modest space chambers cryogenic machines have been employed[9]. The cryopumps are normally operated in conjunction with turbomolecular pumps or ion pumps and with rotary pumps for the initial roughing out. It has been argued that the magnitude of the molecular flux returning to a space vehicle and not the measured pressure, is the important factor in simulating the environment in space. With the high percentage of molecules reflected from the cold walls of the simulation chamber relative to the gas phase molecules, it is then questionable if the degree of vacuum is critical and whether, therefore, the pressure needs to be below 10^{-6} Pa. Barnes and Pinson[10] point out that in any case many of the space simulation experiments are inherently dirty, particularly those involving testing of spacesuits, etc. in what are termed man-rated chambers. A special problem here is the leakage of helium from helium–oxygen breathing apparatus, which is difficult to pump with cryopumps. Perhaps the need for such chambers will be reduced in the future with the setting up of laboratories in space[11].

7.5.2 Surface science

Surface science is probably the area where the inception of ultrahigh vacuum

* Fomblin is the registered trademark of Montedison of Italy

has made its greatest impact. As was seen in Chapter 1 (*Table 1.1*), at pressures around 10^{-4} Pa a surface becomes contaminated with a layer of gas within seconds but if the pressure can be lowered to 10^{-8} Pa, then it would take several hours to build up a gas monolayer. Thus, with an ultrahigh vacuum environment, it has become possible to examine atomically clean surfaces and to develop and monitor surface layers in a controlled fashion. As a result there has been an enormous expansion of surface science. Apart from the studies aimed at a higher degree of understanding of the chemical, structural and electronic properties of surfaces, surface science encompasses a wide variety of applicational studies such as tribology, metallurgy, semiconductor physics, catalysis and thin film technology.

An important outcome of these studies has been the introduction, one might almost say proliferation, of analytical techniques for establishing the structure and composition of the surface or surface layers, each with its own acronym. To give some idea of these techniques, *Table 7.1* lists the best known ones together with the general process and information that is obtainable. It is beyond the scope of this book to enter into duscussion of these techniques and the reader is referred to textbooks on the subject[12,13]. Suffice to say that most of the techniques involve the measurement of the energies, and sometimes species, of charged particles which are derived from the surface and that, in order to ensure the surface remains unchanged during the measurements, ultrahigh vacuum techniques are a necessity.

Because of the wide interest in surface analysis it has become viable to manufacture dedicated analytical equipment in spite of the complexity involved and one or two companies now offer a range of such equipments. To give an example, VG Scientific manufacture a versatile equipment under the name of ESCALAB which allows several surface analytical techniques to be applied to a single specimen chamber. A photograph of the instrument is shown in *Figure 7.10(a)* with a schematic diagram showing the various parts in *Figure 7.10(b)*. The system is constructed of stainless steel with ConFlat flange seals and is bakable. The standard pump system uses a diffusion pump with a liquid nitrogen trap but it could be connected to an ion pump or a turbomolecular pump. A titanium sublimation pump handles the extra gas loads in the specimen preparation area. The system provides several analytical options such as XPS, UPS, AES, EELS and LEED. The exact configuration of the instrument will vary slightly depending on the analytical options chosen. The heart of the instrument is a 150° spherical sector analyser which is combined with a transmission transfer lens for measuring electron energies and densities from the sample. The sample can be manipulated to change position and angle so that various points on the surface can be examined. The system can also be fitted with a scanning electron gun for SEM and SAM investigations.

An important consideration in the study of surfaces is the preparation of the specimen to remove contamination layers, or at least to ensure that the surface being examined is reproducible. There are several methods of preparing the surface. It can be heated by an oven or an electron beam or it can be cleaned by sputtering, i.e. bombarded with inert gas ions. This latter can be carried out in a glow discharge in a low pressure of the inert gas or by a beam of ions in a vacuum. A method of obtaining a clean crystalline surface is to cleave the crystal in the vacuum system. Whatever preparation techniques are used, these

TABLE 7.1. Surface analysis techniques

	Title	*Method*	*Information obtained*
ESCA	Electron spectroscopy for chemical analysis	X-ray ejects electrons from various atomic levels. The electron energy spectrum is measured	Surface composition
XPS	X-ray photoelectron spectroscopy		Surface composition
UPS	Ultraviolet photoemission spectroscopy	Similar to XPS using UV light	Surface composition
SIMS	Secondary ion mass spectroscopy	Ionized surface atoms are ejected by impact ions and subjected to mass analysis	Surface composition
LEIS	Low-energy ion scattering	A beam of inert gas ions is scattered elastically by surface atoms	Surface composition
EID	Electron impact desorption	An electron beam strikes the surface and the energy of ejected ions is measured	Adsorbed species and surface layers
LEED	Low-energy electron diffraction	Back scattering of low energy (10–200 eV) electrons measured	Surface structure
HEED	High-energy electron diffraction	Same as LEED but with energies around 20 keV at glancing incidence	Surface structure
RHEED	Reflection high-energy electron diffraction	Similar to HEED	Surface structure and composition
EELS	(ELS) electron energy loss spectroscopy	Incident electrons are scattered inelastically	Surface energy states and composition
HREELS	High-resolution EELS		Identification of absorbed species
AES	Auger electron spectroscopy	Incident high energy electron ejects inner electron from an atom which is replaced by an outer electron. The released energy is given to a third ejected Auger electron	Surface composition
SEM	Scanning electron-microscope	A beam of electrons is scanned over the surface and scattered electrons focused	Topology of surface structure
SAM	Scanning Auger electron microscope	AES with a scanning electron beam	AES over surface

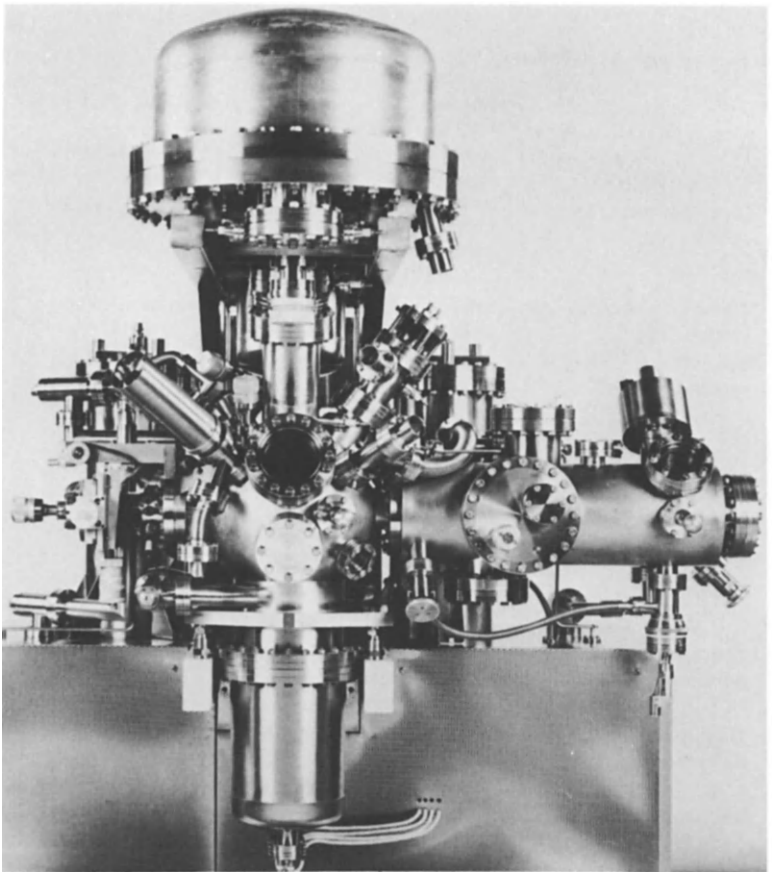


Figure 7.10 (a) Photograph of ESCALAB, a versatile surface analysis apparatus. (Courtesy of VG Scientific)

have to be accommodated in the ultrahigh vacuum system. In the example system above, a preparation chamber is provided between the input vacuum lock and analytical chamber with gate valves for isolation.

7.5.3 Thin film technology

The deposition of thin films for their chemical, physical or electrical properties has found many applications especially in the optics and electronics fields. As a result of the variety of methods available for depositing the thin films, extremely thin layers of very pure materials of almost any composition can be formed on crystalline and amorphous substrates under precisely controlled conditions. Some of these techniques require a vacuum environment whilst others involve a glow discharge at gas pressures in the region of 10–100 Pa. Nevertheless, even in the latter case, it is important that the substrate is clean and that no unwanted reaction takes place between the gas and the layer during formation, particularly if the performance of the final product is

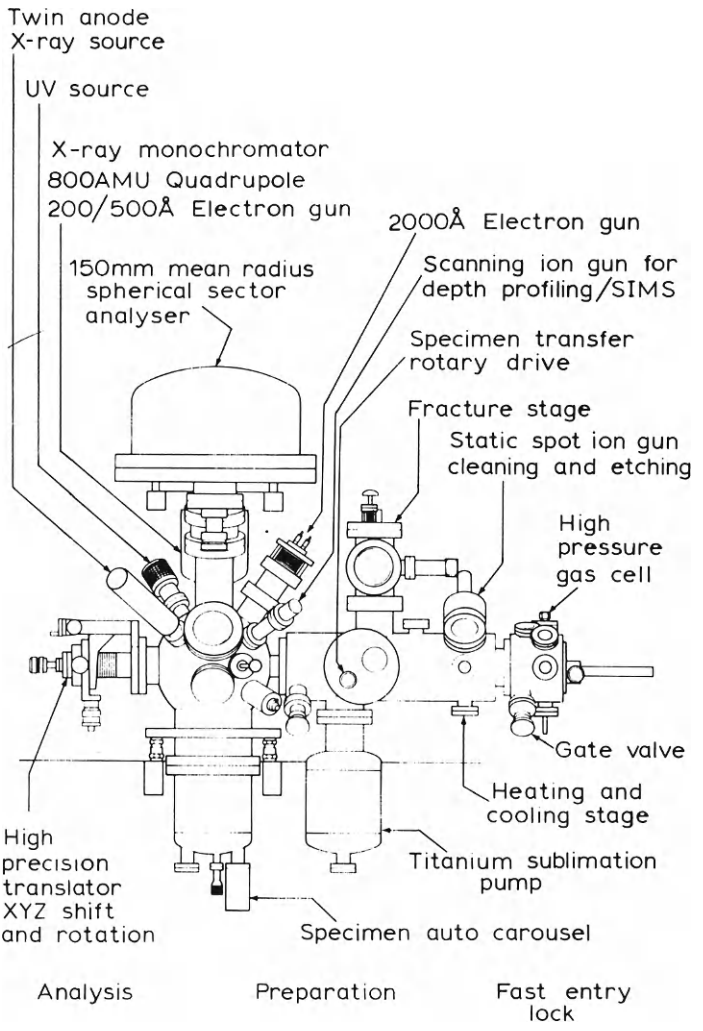


Figure 7.10 (b) Schematic of ESCALAB, detailing the component parts. (Courtesy of VG Scientific)

critical. Thus, for several of the applications the background of residual gases in the system must be of ultrahigh vacuum standard.

The most widely used technique for depositing thin films is by evaporation or sublimation, whereby the material to be deposited is heated, either thermally or by electron beam bombardment, in a vacuum environment. For most of the applications a background pressure below 10^{-4} Pa is sufficient. On the other hand the presence of oil vapour is often disastrous. For some applications the presence of reactive gases even at 10^{-4} Pa can result in imperfections in the deposit or result in non-reproducible magnetic, optic or electric properties. For example, Caswell[14] found that he could not obtain reproducible super-conducting films of tin unless they were deposited in an oil-free system at pressures below 10^{-7} Pa. Similar problems can be encountered

in glow discharge deposition techniques due to residual gas impurities. Most of these techniques employ argon gas to sputter the thin film material on to the substrate, but for some an active gas is introduced to form a compound by reaction with the film material on deposition (reactive sputtering).

Most coating units employ a diffusion/rotary pumping system with suitable trapping to minimize the backstreaming of the pump fluid. However, to meet the more stringent demands placed on some devices, the diffusion pump can be replaced by a cryopump. Several manufacturers offer cryopumped coating units where ultra-cleanliness is required. Apart from freedom from hydrocarbons, the advantage of the cryopump for this application is its high pump speed which allows the system to be pumped down to the ultimate pressure in 10–15 minutes.

Figure 7.11 shows a schematic diagram of the Veeco 7761 cryopumped system which is fairly typical of such units. Although mainly developed for vacuum deposition techniques, the cryopump system can handle the argon gas loads involved in sputtering deposition, albeit at the expense of more frequent regeneration cycles. However, the turbomolecular pump is usually preferred for the glow discharge sputtering units, if an oil-free system is essential. For

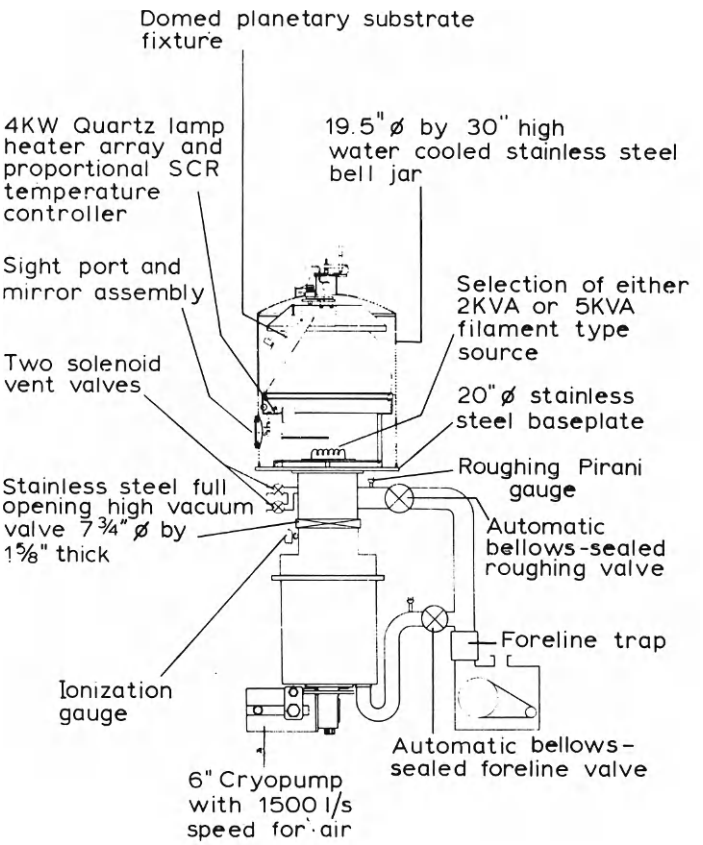


Figure 7.11 Schematic of a typical cryopumped coating unit, the VE-7761 from Veeco Instruments Inc.

some reactive sputtering processes the gases used can be harmful to mechanical pumps in that they can react with the lubricating oil. This is more true of plasma etching systems where halogen containing gases such as CCl_4 and CF_4 are used. The use of synthetic oils, such as perfluoropolyether, for lubrication and chemical filters in the backing line will reduce the effect. It is also advantageous to introduce dry nitrogen instead of air into the gas ballast valve of the rotary pump. Some of the pumps include a combined mechanical and chemical oil filter to reduce the frequency of oil changes and from the health hazard point of view the gases should not be exhausted into the atmosphere without removing the dangerous components.

7.5.4 Semiconductor processing

As integrated circuits become more complex, so the traditional techniques of manufacture are being replaced by methods which involve vacuum processes. This is partly due to the need for higher resolution, partly due to better control on processing and partly economics. Also the vacuum processes tend to be safer from the health hazard point of view. Thus diffusion doping is being replaced by ion-beam implantation, photolithography by electron beam lithography and chemical etching by plasma etching. Most of these processes do not require ultrahigh vacuum and pressures around 10^{-4} Pa are normally adequate. On the other hand, in the research and development of these processes, ultrahigh vacuum systems are often used, if for no other reason than to reduce the number of variables.

One requirement, which chiefly applies at present to semiconductors other than silicon devices, is the deposition of epitaxial layers. Materials that are difficult to produce as reasonable sized crystals in bulk are deposited on a substrate with a similar crystalline structure, to give a large defect-free oriented crystal layer. As an example, cadmium mercury telluride can be grown on cadmium telluride for infra-red photodetectors. Epitaxial layers can be grown by liquid or vapour interaction with the solid substrate at elevated temperatures. However, a new method, requiring an ultrahigh vacuum environment, is now becoming established and is known as molecular beam epitaxy (MBE)[15]. MBE is a controlled form of evaporation in which atomic or molecular beams are produced in small furnaces having an orifice (Knudsen effusion sources) and are condensed on to a heated substrate in vacuum. A characteristic feature of MBE is its ability to provide much smoother surfaces and interfaces than alternative growth techniques, which makes it of particular interest where optical properties are exploited such as in solid state lasers. It also offers better control on doping profiles and composition and allows *in situ* preparation of the substrate.

Figure 7.12 shows a schematic diagram of an MBE chamber, constructed in the author's laboratory, for depositing and monitoring the epitaxial layers. More than one source is usually required to give the correct composition of the layer. The surface can be monitored by an electron beam which can be scanned over the surface to monitor thickness and structure. *Figure 7.13* shows the complete ultrahigh vacuum system for the MBE equipment. This particular system is pumped by a diffusion pump with a liquid nitrogen trap. There is also a titanium sublimation pump which can be brought into operation when dealing with high gas loads and for use in a sealed-off situation. The whole

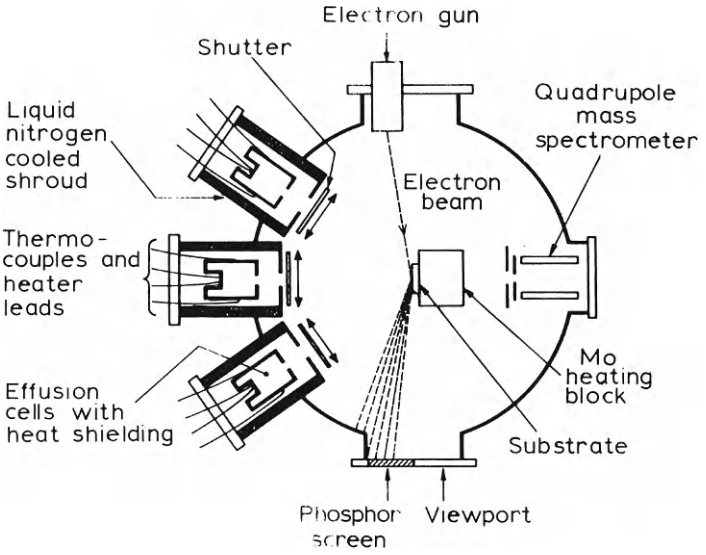


Figure 7.12 Schematic diagram of an MBE chamber

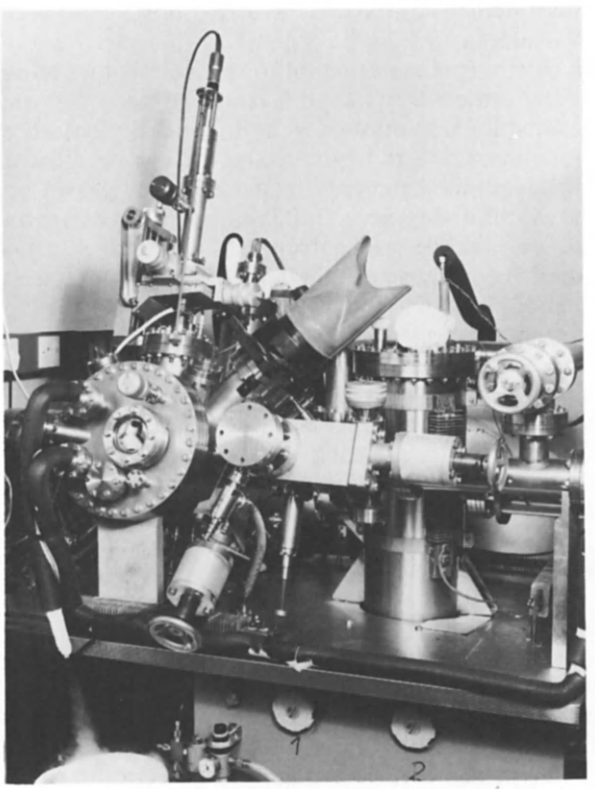


Figure 7.13 MBE facility at Philips Research Laboratories

system above the diffusion pump can be baked up to 200°C, limited by the Viton A seals of the gate valves employed for isolating the system from the pump and for the sample entry lock. After bakeout, base pressures of the order of 10^{-8} Pa are regularly achieved. A review article on the technique and applications has been written by the scientists mainly responsible for the MBE programme at Philips[16] to which the readers may like to refer.

The vapours used in the processing of semiconductors for integrated circuits are often toxic or corrosive. For example, arsenic compounds are used for some ion implantations and carbon tetrafluoride is a common etchant for silicon oxides. Special care has to be taken in dealing with such hazardous gases and the precautions mentioned in the last section must be adhered to, namely a closed pumping system and special mechanical pumps. Contamination of component parts, i.e. deposition of arsenic on the electrodes of an ion-implanter, necessitates the dismantling of the equipment for cleaning. It is important that such processes are carried out by a competent person who has been trained in the methods to be used and is fully aware of the dangers.

7.5.5 Other applications

There are several other applications, mainly in the research and development area, where an ultrahigh vacuum environment is advantageous if not a necessity. For example there is a wide field where mass spectrometry is used as a tool; in the medical field for breath analysis, in the petrochemical industry and even in the iron and steel industry, where flue gases from blast furnaces are analysed. In some of these applications small amounts of impurities are being looked for and a low base pressure is then essential.

Although electronics is dominated by solid state devices, there are still some areas where the electronic vacuum tube has not been replaced. Semiconductors cannot match the power handling capabilities of some of the microwave and transmitting valve types and for large high-resolution displays, the cathode-ray tube reigns supreme in spite of considerable effort to find an alternative. There is continuing investigation of such devices in which ultrahigh vacuum plays an important role. As an illustration, *Figure 7.14* shows an ion pumped ultrahigh vacuum system used for investigating the design of a flat cathode-ray tube proposed by Lamport *et al.*[17]. Vacuum tubes are also used for photon detection and imaging. In particular image intensifiers have been used extensively for low light-level viewing both for military and civilian applications. The so-called third generation image intensifier exploits caesiated GaAs as the photocathode, where sensitivities of the order of 2 mA lm^{-1} have been quoted, which are at least an order higher than those for the multialkali photocathodes previously used. Such sensitivities can be reached only if the GaAs surface is very clean and free from any contamination and this requires manufacture in an ultrahigh vacuum environment.

Outside the electronics field, ultrahigh vacuum is used in a variety of research studies such as the understanding of vacuum breakdown, atomic and molecular collision processes, isotope separation, sputter mechanisms and the examination of both inorganic and organic materials. There is also a fair amount of research and development effort being devoted to the physics and technology of vacuum production and measurement itself.

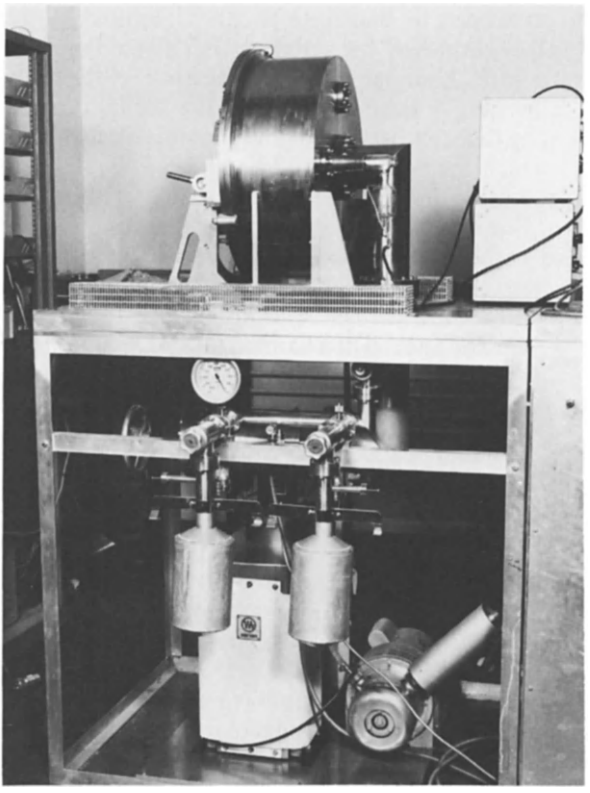


Figure 7.14 An ion pumped system for investigating flat cathode ray tubes. Lamport *et al.*[17]

In summary, the advancement of vacuum technology, which started with the quest for better vacuum in the 1950s, has had significant impact on research and development laboratories over a wide field and in some industries has dramatically changed manufacturing processes. Although the advancement in vacuum technology has been less dramatic since the 1960s, it has by no means petered out and no doubt the further understanding of vacuum physics and related subjects, such as surface physics, will result in new applications developing in the future.

7.6 References

1. DENNIS, N. T. M., COLWELL, B. H., LAURENSON, L. and NEWTON, J. R. H., *Vacuum*, **28**, 551, (1978)
2. HENNING, J., *Vacuum*, **21**, 523, (1971)
3. HAEFER, R. A., *Vacuum*, **30**, 193, (1980)
4. KUBIAK, R. A. A., LEONG, W. Y., KING, R. M. and PARKER, E. H. C., *J. Vac. Sci. Technol.*, **A1**, 1872, (1983)
5. LUCAS, J., GRIFFITHS, I. and GOODWIN, C., *Vacuum*, **30**, 159, (1980)
6. TRICKETT, B. A., *Vacuum*, **28**, 471, (1978)
7. GOETZ, D. G., *Vacuum*, **32**, 703, (1982)

8. BUDGEN, L. J., *Vacuum*, **32**, 627, (1982)
9. FORTH, H. J. and FRANK, R., *Proc. 7th Internat. Vacuum Cong., Vienna*, 61, (1977)
10. BARNES, C. B. and PINSON, J., *Proc. 4th Internat. Vacuum Cong.*, Inst. of Physics Conf. Series No. 5, 219, (1968)
11. KLEBER, P. K., *Proc. 7th Internat. Vacuum Cong. Vienna*, 333, (1977)
12. GZANDERNA, A. W. (Editor), *Methods of Surface Analysis*, Elsevier Scientific, (1975)
13. BACH, I. (Editor), *Electron Spectroscopy for Surface Analysis*, Springer-Verlag, (1977)
14. CASWELL, H. L., *J. Appl. Phys.*, **32**, 105, (1961)
15. CHO, A.Y. and ARTHUR, J. R., *Progress in Solid State Chemistry*, **10**, edited by J. McCALDIN and G. SOMERJAI, 157, (1975)
16. FOXON, C. T. and JOYCE, B. A., *Current Topics in Materials Science*, 7, edited by E. KALDIS, North Holland Publishing Co., Chapter 1, (1981)
17. LAMPORT, D. L., WOODHEAD, A. W., WASHINGTON, D. and OVERALL, C. D., *IEE Proc.*, **131 Pt. I**, 10, (1984)

Chapter 8

Leak detection

8.1 Basic techniques

Gas leaks in a system can be a vacuum engineer's nightmare, difficult to locate and, when found, sometimes difficult to cure. In ultrahigh vacuum systems, leaks which may be extremely small will still be important and, because of their size, one may suspect that they will be more difficult to trace. In practice, however, this is not necessarily the case. The lower base pressure and the high sensitivity of the gauges used for ultrahigh vacua, generally allow smaller leaks to be detected and location is often easier than for a higher pressure system.

The presence of a leak is manifest in the inability to attain the predicted ultimate vacuum pressure. To ascertain whether the poor vacuum is due to a leak or to outgassing, it is useful to isolate the system from the pump or to render the pump inoperative and to monitor the rise in gas pressure. If the pressure rises at first rapidly but tends to level out at a higher value, then outgassing or contamination by a material of high vapour pressure could be the cause. On the other hand, if the pressure continues to rise without showing a saturation effect then a leak can be suspected. If the system is such that various parts can be isolated with valves, it may be possible by sequentially opening and closing valves, to locate the general area of the leak. For small leaks, the interval of time required to reach a decision could be inordinately long and, if the expected ultimate pressure cannot be attained, it is often judicious to check for a leak first before looking for other causes of gas influx.

The basic method for detecting small leaks in the vacuum envelope is to surround the envelope with a plastic bag and introduce a gas which displaces the air and will then flow through the leak. The test gas is detected by a gas pressure sensor within the system, whose sensitivity is dependent on the gas species and whose response will thus change as the test gas infiltrates. As pointed out in Chapter 4, the sensitivity of an ion gauge depends on the gas composition and, in particular, in a Bayard–Alpert gauge (BAG) the sensitivity for helium is only one-fifth of that for nitrogen (*Table 4.1*). Thus the BAG, which is incorporated in most ultrahigh vacuum systems, provides a convenient sensor for leak detection when using helium as the test gas.

It is also possible to exploit the ion pump, where this is employed, to detect

leaks by measuring the change in pump current when the gas changes composition. Neither of these sensing techniques, however, may be sensitive enough for extremely small leaks and one must then resort to special sensors.

Such a sensor is the mass spectrometer specifically designed for leak detection, in which the head is tuned to give maximum sensitivity to the test gas. The advantage of the high sensitivity of mass spectrometer leak detectors has resulted in their universal use for finding small leaks in vacuum systems and most vacuum equipment manufacturers supply leak detectors which incorporate the mass spectrometer head with valves, pumps and electronic output. The equipment can either be used as a stand-alone unit for testing vacuum components or transported to the vacuum system and coupled in for testing *in situ*.

Having established the presence of a leak the next procedure is to locate it. This can be carried out by probing the exterior of the system with a jet of the test gas. The problem here is that the jet may only impinge on the leak for a short period whilst the jet is being moved over the system. Because of the time taken for the gas to travel through the leak to the detector, the presence of the probe gas may not be detected until the probe has passed the leak. The search can therefore be a protracted process. Pre-testing of the component parts of the vacuum envelope and intelligent guessing of the likely area of the leak based on experience, can speed up the process. The sensitivity of leak detection can be enhanced by allowing accumulation of the probe gas, for example by valving off from the pump that part of the system associated with the search area.

If the leak is large, so that the residual pressure is high, the effect of the leak on the total current may not be noticed, especially as the probe gas will be considerably diluted by the air. The use of acetone in these circumstances may be a better solution. It could change the pressure significantly by momentarily plugging the leak or by evaporating.

Curing the leak, once it has been found, may not be straightforward, it will depend on the cause. If it occurs at a demountable seal, tightening the bolts or replacing the gasket may be all that is required to eliminate the leak. If it is due to faulty material, to a bad weld or to a damaged seal, for example at a metal–ceramic interface, then the only permanent answer is to replace the faulty component. There are some sealant liquids which can be applied to small leaks to seal them effectively by diffusing into the hole and solidifying; these were mentioned in Section 2.5. Although the resultant seal may be bakable at low temperatures, they should only be considered as a temporary remedy until the component can be replaced at a more convenient time. If the leak is due to a weld or a glaze seal then the component may be repaired off the system by rewelding or reglazing respectively.

8.2 Performance requirements

In Chapter 7 an example of a system was considered, of volume $0.028\ m^3$, pumped by a diffusion pump with an effective pump speed of $0.25\ m^3\ s^{-1}$. The relation between the ultimate pressure and the influx of gas for this system is given by

$$0.25\,P_{ult} = Q_i \text{ Pa m}^3\text{s}^{-1} \tag{8.1}$$

Thus, if an ultimate pressure of 10^{-10} Pa is to be achieved on a continuously pumped system, then the leak rate must be less than Q_i and therefore less than 2.5×10^{-11} Pa m^3s^{-1}. For an ultimate pressure of 10^{-8} Pa the corresponding leak rate must be less than 2.5×10^{-9} Pa m^3s^{-1} and so on.

The rise of pressure with time, assuming the pump is valved off and that the influx gas is mainly due to a leak, is given by

$$\frac{dP}{dt} = \frac{Q_i}{V} \tag{8.2}$$

which for a leak rate of 2.5×10^{-11} Pa m^3s^{-1} would give a rise in pressure of 8.9×10^{-10} Pa within a second and 3.2×10^{-6} Pa in an hour for the example system. Thus, a much lower leak rate will be required for a system which is to be isolated from the pump than can be tolerated for a continuously pumped system.

For the leak detector to be effective it must be capable of measuring leak rates below these values. To determine the lowest leak rate that can be detected, one needs to know the effect that the gas change has on the detector reading when the test gas is introduced and how this is related to the size of the leak. Two cases can be considered, one where the leak rate is fairly high, above 10^{-6} Pa m^3s^{-1}, and where viscous flow conditions are likely to apply to the gas flow through the leak and the other where the leak rate is lower and molecular flow conditions apply.

In the case of viscous flow conditions the leak rate for air Q_{La} will be given by

$$Q_{La} = \frac{B}{\eta_a}(P_A^2 - P_V^2) \tag{8.3}$$

where η_a is the viscosity for air, P_A is atmospheric pressure, P_V is the pressure in the vacuum chamber and B is a constant which depends on the geometrical dimensions of the leak. We can equate Q_{La} to the effective pump speed for air, S_a as

$$Q_{La} = \frac{B}{\eta_a} P_A^2 = S_a P_V \tag{8.4}$$

In this equation it is assumed $P_A \gg P_V$.

Assuming that in our leak test the air is completely replaced by the search gas, i.e. by enclosing the system in a gas bag, then the corresponding equation for the search gas can be written as

$$Q_{Ls} = \frac{B}{\eta_s} P_A^2 = S_s P'_V \tag{8.5}$$

where P'_V is the new pressure in the vacuum chamber. Thus, the change in pressure in the chamber due to changing the air leaking through with a test gas is given by

$$\Delta P = P'_V - P_V = BP_A^2\left(\frac{1}{\eta_s S_s} - \frac{1}{\eta_a S_a}\right) \tag{8.6}$$

Equation (8.6) can be rewritten as

$$\Delta P = \frac{Q_{La}}{S_a}\left(\frac{\eta_a S_a}{\eta_s S_s} - 1\right) \tag{8.7}$$

For a detector which has a sensitivity of K_a for air and K_s for the search gas, the change in detector reading ΔG will be given by

$$\begin{aligned}\Delta G &= K_s P'_v - K_a P_v \\ &= BP_A^2\left(\frac{K_s}{\eta_s S_s} - \frac{K_a}{\eta_a S_a}\right)\end{aligned} \tag{8.8}$$

A similar equation can be derived for the lower leak rate where molecular flow conditions apply. From Section 1.3.4 it was seen that the flow rate under these conditions depends on the pressure difference across the leak and the square root of the mass of the gas molecules. This gives the change in gauge reading due to replacing the air with a test gas as

$$\Delta G = CP_A\left(\frac{K_s}{S_s\sqrt{m_s}} - \frac{K_a}{S_a\sqrt{m_a}}\right) \tag{8.9}$$

where C is a constant which depends on the geometrical dimensions of the leak. It is related to the conductance for air, C_a, by the relationship $C = C_a\sqrt{m_a}$.

Consider first the case of a large leak of $Q_{La} = 10^{-6}$ Pa m^{3}s^{-1}. Equation (8.7) then gives

$$\Delta P = \frac{10^{-6}}{S_a}\left(\frac{\eta_a S_a}{\eta_s S_s} - 1\right) \tag{8.10}$$

If helium is used as the test gas then its viscosity is very near to that of air so that the function in the bracket will depend on the pump speed for the two gases alone. In the case of a diffusion pump the speed for different gases is roughly dependent on the reciprocal of the square root of the mass of the gas molecules and therefore such a pumped system will give a drop in pressure of

$$\Delta P = \frac{10^{-6}}{S_a}\left(1 - \frac{\sqrt{m_{He}}}{\sqrt{m_{N_2}}}\right) = \frac{6.2 \times 10^{-7}}{S_a}\ \text{Pa} \tag{8.11}$$

If we take our example with $S_a = 0.25$ m^{3}s^{-1} then the drop in pressure will be 2.5×10^{-6} Pa. Since for the BAG the sensitivity for helium is lower than for air, the apparent drop in pressure will be even greater and the leak will be easily detected especially if the original reading is 'backed-off' by applying a reverse voltage to zero the meter (see Section 8.3).

The more interesting case is for the small leak where Equation (8.9) applies. Again if we consider a diffusion pump system where the pump speed is dependent on m$^{-1/2}$ so that $S_s m_s^{1/2} = S_a m_a^{1/2}$ then $P_v = P'_v$ and

$$\Delta G = \frac{Q_{La}}{S_a}(K_s - K_a) \tag{8.12}$$

since $Q_{La} = C_a P_A = CP_A/\sqrt{m_a}$ when $P_A \gg P_v$.

For helium using a BAG the relative sensitivities are given in *Table 4.1*. This results in $K_s - K_a = -0.8K_a$ and, therefore, ΔG in terms of the equivalent

pressure drop with a 0.25 m^3s^{-1} pump is

$$\Delta G(\text{Pa}) = \frac{-\Delta G}{K_a} = 3.2\, Q_{\text{La}} \tag{8.13}$$

Since the minimum pressure for the BAG is around 10^{-8} Pa on a full scale of say 10^{-7} Pa and, as will be shown later, a 0.1% change in the full scale reading can be measured, then the minimum detectable leak rate for this system would be 3×10^{-11} Pa m^3s^{-1}.

The above equations assume that the atmosphere around the vacuum system is completely replaced by the test gas, which has been taken as helium in the example. This situation appertains to the case where the system is surrounded by a gas bag filled with the test gas. When the system is being probed with a gas jet the exterior gas will be a mixture of air and the test gas and the equations are modified. If we assume that the concentration of the test gas is x atmospheres where $x < 1$ and that η_x is the viscosity of the mixture, then for this mixture it follows that

$$\frac{BP_A^2}{\eta_x}(1-x) = S_a P'_a$$

and (8.14)

$$\frac{BP_A^2}{\eta_x} x = S_a P'_s$$

where P'_a is the partial pressure of air in the system and P'_s is the partial pressure of the test gas. The total pressure in the system

$$P'_v = P'_a + P'_s = \frac{BP_A^2}{\eta_x}\left(\frac{1-x}{S_a} + \frac{x}{S_s}\right) \tag{8.15}$$

and therefore

$$\Delta P = P'_v - P_v = BP_A^2 \left\{\left(\frac{1-x}{S_a} + \frac{x}{S_s}\right)\frac{1}{\eta_x} - \frac{1}{\eta_a S_a}\right\} \tag{8.16}$$

which can be compared with Equation (8.6). If η_x is close to the viscosity for air then we see that the pressure difference is reduced by the fraction x from the case where the total gas environment has changed. However, in the probe gas case the transient effects must also be considered. Apart from the delay in gas penetrating the leak as the probe comes over the hole, there will also be a delay before the equilibrium in the gas ratios is reached within the system, as the air is pumped away and the test gas pressure builds up. If the pumping equation (7.4) is taken, i.e.

$$P_t = \frac{Q_L}{S} - \left(\frac{Q_L}{S} - P_0\right)\exp\left(-\frac{S}{V}t\right) \tag{8.17}$$

then for the test gas, P_0 is 0 at $t = 0$ so that the build up in test gas partial pressure is given by

$$P_{ts} = \frac{Q_{LS}}{S_s}\left[1 - \exp\left(-\frac{S_s t}{V}\right)\right] \tag{8.18}$$

On the other hand the partial pressure of the air starts at the ultimate Q_{La}/S_a and falls exponentially

$$P_{ta}=\frac{Q_{La}}{S_a}\exp\left(\frac{-S_a t}{V}\right) \tag{8.19}$$

Thus the partial pressure of the air will fall with a time constant V/S_a whilst the partial pressure of the test gas will rise with a time constant of V/S_s. The total pressure at time t will be

$$P_{tT}=\frac{Q_{LS}}{S_s}\left(1-\exp\left(\frac{-S_s t}{V}\right)\right)-\frac{Q_{LA}}{S_a}\exp\left(\frac{-S_a t}{V}\right) \tag{8.20}$$

When the probe gas moves away from the leak the equations are reversed, the partial pressure of the test gas declines whilst the air partial pressure increases. This is dealt with in more detail by Robinson[1]. The interesting point to note from Equation (8.20) is that if the pump speed for the test gas is very much less than that for air, then a fast response will occur. This is because the faster pump speed for air will rapidly pump it away whilst the slow speed for the test gas allows the test gas pressure to build up quickly. On the other hand when the jet of test gas moves away from the leak the partial pressure of the test gas will reduce very slowly and because of the higher pump speed for air the partial pressure of the air will take longer to build up. As a result a much longer recovery time will occur.

The importance of pump speed in the detection of a leak is evident in Equations (8.8) and (8.9). Lowering the pump speed is one way of improving the sensitivity of the detection system. However, there are other ways, one of which is to allow a build up of pressure before taking readings on the detector. The system is pumped down to the lowest pressure attainable and then the pump is closed off. After, say, a minute the pressure is read. The pump is reconnected and the test gas applied. After a suitable time to allow the air to be replaced by the test gas, the pump is again closed off and after one minute the new pressure reading is taken. The change in gauge reading will be given by

$$\Delta G=\frac{60}{V}CP_A\left(\frac{K_s}{S_s\sqrt{m_s}}-\frac{K_a}{S_a\sqrt{m_a}}\right) \tag{8.21}$$

where V is the closed-off volume.

This is a rather simplified representation since the outgassing of the system is ignored. However, it does suggest that for our example system of $V=0.28\ \text{m}^3$, leak rates of 10^{-12} Pa m^3s^{-1} might be detected with a BAG. Certainly where a mass spectrometer head is employed and therefore outgassing is less of a problem, significant improvements in sensitivity have been demonstrated[2] using this technique.

Another method of improving sensitivity, which has been used in the past but which is of less interest today, has been to interpose a selective gas filter between the system and the detector which will allow the test gas to pass through alone. An example of such an arrangement is to use a heated palladium tube between the ion gauge and the system with hydrogen as the test gas. As pointed out in Section 2.3.3, the permeation rate for hydrogen through palladium is at least two orders higher than through other metals, whilst the permeability to other gases is negligible. Since, however, this method requires a

dedicated detector and filter for leak detection, it has been superceded by the mass spectrometer leak detector (see Section 8.6).

8.3 Leak detecting with an ion gauge

This method is particularly commended for ultrahigh vacuum systems, since at least one sensitive ion gauge will normally be fitted. It offers a cheap and simple method of detecting leaks, limited only by the minimum pressure reading of the gauge. This implies that if the required ultimate pressure in a continuously pumped system is within the range of the gauge then, in principle, any leak which would cause the pressure to rise above this required value is capable of detection by the gauge. For many applications, therefore, the performance that can be attained by this method would be adequate.

To obtain the highest sensitivity, a test gas should be chosen which will give the greatest difference in gauge reading when substituted for the air environment. As pointed out in the previous section this depends not only on the sensitivity of the gauge to the two different gases but also on the speed with which they are pumped by the system. This in turn will depend on the type of pump that is used. It has already been shown that for small leaks, where molecular flow conditions through the leak are likely to prevail, there is no change in pressure when pumping with a diffusion pump, whatever test gas is used (cf. Equation (8.12)). This is because the leak rate and the pumping speed are both related to the square root of the gas mass to the first approximation, i.e.

$$P_V = \frac{Q_{La}}{S_a} = \frac{Q_{Ls}}{S_s} \tag{8.22}$$

Because there is no change in actual pressure when the test gas is substituted, the test gas for diffusion pumped systems is chosen to give the greatest change in gauge sensitivity and hence in the apparent indicated pressure. From *Table 4.1* it is seen that if a Bayard–Alpert gauge is used, then helium is a satisfactory test gas. However, other gases can be used. In particular, Calor gas (butane), with a sensitivity relative to nitrogen of 4.46, was found by Blears and Leck[3] to give the best results out of the several gases they tested. For larger leaks the relative pump speed of the gases becomes important, for although the viscosity of a gas depends on the square root of the mass, it also depends on the molecular diameter, δ. The larger the diameter the lower the viscosity, since

$$\eta = \frac{5}{16\delta^2}\left(\frac{kmT}{\pi}\right)^{1/2} \tag{8.23}$$

As a result, the viscosity does not vary as markedly with mass as does the pump speed, so that the relative change in gauge reading, $\Delta G/P$, is greater for larger leaks than for the small leak because the increase due to the sensitivity of the gauge is enhanced by the change in pump speed.

For a different type of pumping system a different conclusion will be reached. With an ion pumped system the significant factor is the low pump speed for inert gases compared to that for active gases. This implies that if helium is used as the test gas, the lower pump speed will result in an increase in

actual pressure, counteracting the lower gauge factor for helium. Therefore, it is not a good test gas to use in such a system. On the other hand, with the heavier inert gas, argon, for which the sensitivity is 1.5 times that of nitrogen, the low pump speed and resulting rise in pressure on applying the test gas is augmented by the increase in sensitivity. The pumping speed for argon is at best only 25% of that for nitrogen and thus for a small leak, where Equation (8.9) applies, it is seen that

$$\frac{K_s}{S_s\sqrt{m_s}} = \frac{1.5}{0.25}\cdot\frac{\sqrt{28}}{\sqrt{40}}\frac{K_a}{S_a\sqrt{m_a}}$$

$$= 5\frac{K_a}{S_a\sqrt{m_a}} \tag{8.24}$$

and from Equation (8.9) it is seen that ΔG is relatively large. Therefore, for the ion pumped system argon is a good test gas to use.

For a turbomolecular pumped system, the pump speed is not dependent on the gas species. However, the compression ratio is very dependent on the gas species, varying exponentially with $m^{1/2}$. For example, one manufacturer's data give the compression ratio for nitrogen as 8×10^8 whereas for helium it is only 2.5×10^4. This would mean that, for a backing pressure of 1 Pa, the pump can pump nitrogen down to a pressure of $\sim 10^{-9}$ Pa but will not be able to pump a system filled with helium to a pressure below 5×10^{-5} Pa. If the leak is such that the pressure is well below 5×10^{-5} Pa, then helium passing through the leak would raise the pressure significantly and give a large change in gauge reading in spite of the lower gauge sensitivity. It would not be as satisfactory, however, if the turbomolecular pump is operated with a lower backing pressure, since then the rise in pressure will be less and might be counteracted by the lower sensitivity.

In a cryopumped system the effect of helium will also be to increase the pressure, since it will not be pumped so efficiently by the cryopanel. The problem here could be in pumping the helium away after having discovered the presence of a leak. Often in cryopumped systems it is worth trying other gases such as CO_2 and oxygen as well as the inert gases.

The change in gauge reading on the normal gauge control box may be small relative to the actual pressure reading, particularly when using a gas search probe. In order to amplify the signal due to the reading change, it is necessary to back-off the base pressure value, i.e. to zero the meter reading by applying a signal equal and opposite to the steady value. This is rarely possible on the normal gauge control box. However, since the output is stabilized, the voltage across the meter can be tapped off and compared against a variable reference voltage. The voltage difference can then be amplified and fed into a meter having a central zero. With air surrounding the system, the reference voltage is adjusted to give a zero reading. When the test gas is applied the meter will read either negative or positive depending on whether a lighter or heavier gas is used. In practice, auxiliary units are available on the market for leak detection with an ion gauge, in which rather more sophisticated circuits are employed that respond to changes in signal without the need for a backing-off voltage. Sensitivities giving a full-scale reading for a 0.1% change in the normal full-scale gauge reading can be obtained, although there will be a minimum rate of

change to activate the circuit. With such a control box and a BAG, detection of leaks down to 10^{-12} Pa $m^3 s^{-1}$ can be achieved under optimum conditions.

8.4 Leak detecting with an ion pump

The equilibrium current drawn by an ion pump is proportional to the pressure but, because it is an ionization current, its absolute value depends upon the composition of the gas within the pump. Thus changing the gas composition by passing a test gas through a leak will change the equilibrium current and, in a similar way to the ion gauge, the ion pump can be used as a sensor for leak testing. A change in voltage applied to the ion pump will also give an associated change in pump current and it is therefore necessary to stabilize the pump power supply voltage when using this method of leak detection. This can be done by replacing the normal power supply with a stabilized supply incorporating a current change indicator but this would be an expensive solution. A better method is to provide voltage stabilization and the current change detector in a separate portable unit which can be used in conjunction with the normal unstabilized power supply when leak checking is required. A simplified block diagram of such an arrangement is shown in *Figure 8.1*. Because both the ion pump and the power supply are separately earthed, the current has to be measured in the HT line. For reasons of safety, however, the current change meter should be situated near earth potential. One method of overcoming this conflict is by an optical link, using a photoemitter in the HT circuit and a photodetector in the meter circuit.

The change in the gas not only alters the current due to the different ionization coefficients but also affects the pumping speed of the ion pump, which in turn will change the pressure within the pump. The result is a further change in current. Using argon as the test gas increases the current, whilst oxygen will decrease it. The effect and the sensitivity are similar to those experienced with an ion gauge used for leak detection in an ion pumped system, as described in the previous section.

8.5 The halogen leak detector

The halogen leak detector is a dedicated system with the sensor designed specifically for leak detection. Although not as sensitive as the ion gauge under optimum conditions, the minimum detectable leak is around 10^{-9} Pa $m^3 s^{-1}$, it is an economic method which is useful for systems working in the upper ultrahigh vacuum pressure region.

The detector is based on the phenomenon of surface ionization[4]. If gas molecules impinge on an incandescent metal surface they may evaporate partly as neutral molecules and partly as positive ions. The ratio of the number of ions leaving the surface per second, N_+, to the number of molecules leaving per second, N_0, is given by the Sara–Langmuir relation which can be expressed as

$$\frac{N_+}{N_0} = \beta \exp\left\{\frac{-eV_1 + \phi}{kT}\right\} \tag{8.25}$$

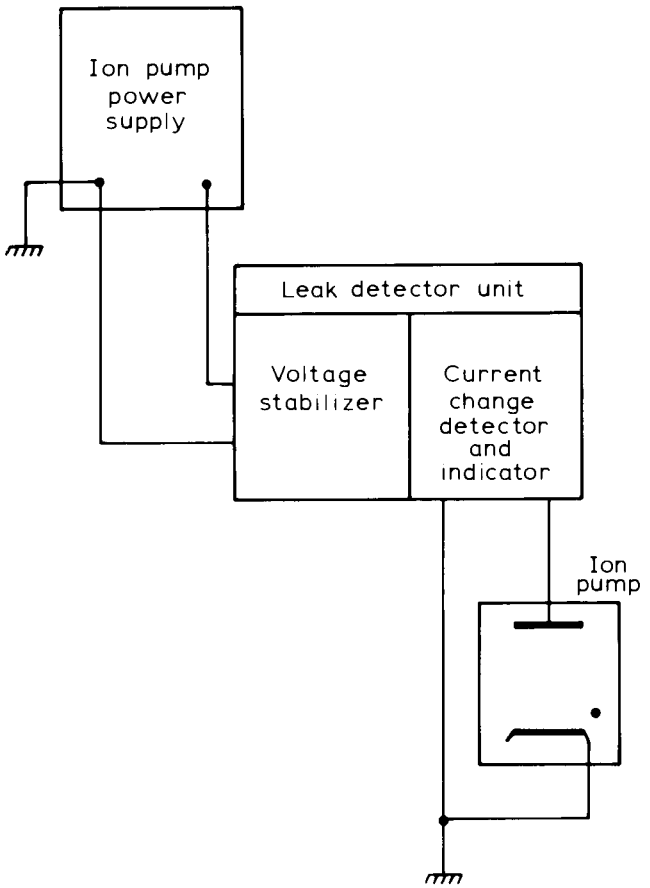


Figure 8.1 Arrangement for detecting leaks with an ion pump

where β is a constant depending on the gas and metal, ϕ is the work function of the metal surface, e is the electronic charge and V_I the ionization potential of the gas.

In the halogen leak detector, the ion current from a heated platinum electrode is measured when a halogenated gas such as Freon-12 (CCl_2F_2) is leaked into the system. The ionization potential of the halogenated gases is around 9–10 V as opposed to 15 V for nitrogen. Since the work function of platinum is around 5.4 eV, changing from nitrogen to Freon-12 more or less squares the exponential term of Equation (8.25). The ion current increases with temperature but, since the background ion current increases as well as the signal current, there is in fact an optimum temperature for the best signal-to-noise ratio, which is around 800°C.

In one practical device, the head consists of an indirectly heated platinum cylinder with a second electrode around it which is negatively biased to collect the ions[5]. The electrodes are contained in a metal envelope which can be connected to the vacuum system via a suitable demountable seal. Using the system pumps, the leak is searched with a Freon gas probe. The principle will also work at atmospheric pressure and a sensor can be designed to act as a

sniffer-head outside the system which is then filled to a pressure above atmospheric with an air–Freon gas mixture. This method is often used for testing components before assembly. Halogen leak detectors as complete units with a head and drive circuit are commercially available.

8.6 The mass spectrometer leak detector

The most sensitive and commonly used dedicated leak detection system for high vacuum application employs a mass spectrometer head which is tuned to the test gas. Almost any mass spectrometer can be adapted, but in practice most commercial leak detecting equipment of the mass spectrometer type employs a magnetic sector analysing head.

To obtain the maximum sensitivity the resolution is sacrificed to some extent. It is important, therefore, to choose a test gas whose mass number is unique, which can be easily resolved and which is not present to any extent in the atmosphere. Helium, with a mass number of 4, fulfils all these conditions and being also inert is ideal. Because helium is universally exploited, the mass spectrometer leak detector is often referred to as a helium leak detector.

Most helium leak detectors use a 90° or 180° magnetic sector-type analyser with a Nier ion source and a conventional ion collector, as described in Chapter 4. The slit widths are chosen to give the best compromise between sensitivity and resolution. Since a resolution not much above four is required, the slit widths and therefore the ion current and sensitivity can be much greater than for the residual gas analyser. However, although the residual gases such as nitrogen, oxygen and water vapour do not have mass numbers close to mass four, they often cause a background signal. This is because some of these ions will be scattered or will bounce off the surfaces within the analyser to produce a change in velocity or direction which allows them to pass through the exit slit and be collected. The percentage of the residual gas ions which will be collected is very small, but since the total residual gas ion current will be large compared to the helium current, the background can be the limiting factor to the sensitivity. A design of analyser aimed at reducing the background current has been described by Warmoltz[5]. The design basically consisted of two 60° magnetic sector analysers in series, in the arrangement illustrated in *Figure 8.2*. From the first sector the helium ions plus the background ions pass through the exit slit into the second sector, where the background ions are considerably reduced. The original analyser was developed for laboratory use but it has since been taken up commercially. The Veeco range of mass spectrometer leak detectors incorporate analyser heads based on the Warmoltz design. Whatever analyser head is used, it is important that the head itself does not produce background noise as a result of gas contamination. To ensure adequate outgassing the head, and indeed the rest of the system, should be readily bakable. Some of the commercial leak detectors hold the entire spectrometer head at an elevated temperature at all times to prevent gas adsorption, whilst others have a continuously heated ion source.

Unlike the other leak detection methods, the mass spectrometer leak detector is presented as a complete system. This normally includes the spectrometer head and the vacuum system consisting of ultrahigh vacuum pump, backing pump, vacuum gauges and valves and the electronics necessary

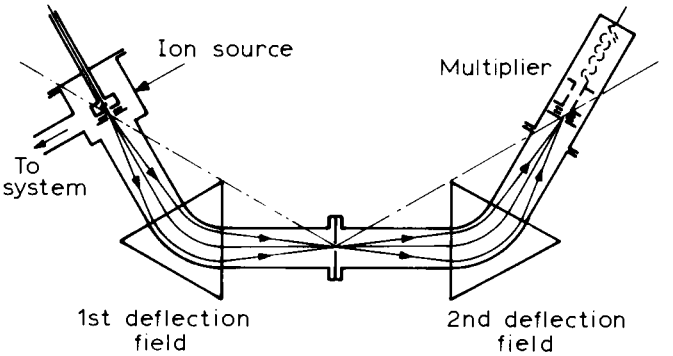


Figure 8.2 Schematic of the double magnetic sector analyser according to Warmoltz[5]

to control the system. A typical leak detection unit is shown diagrammatically in *Figure 8.3*. Most units are versatile and can be applied either to leak testing of components or can be connected up to the system in which a leak is suspected. Because it has its own pumping system, the pump speed can be controlled, and as a result, the spectrometer can be calibrated to measure directly the leak rate (see Section 8.7). The pump speed also influences such parameters as response time and background signal.

When used for leak testing an ultrahigh vacuum system, the mass

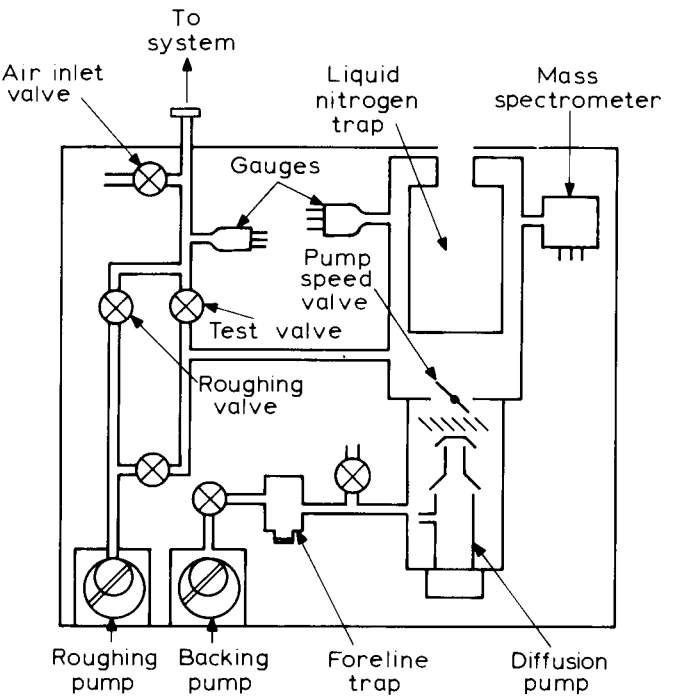


Figure 8.3 Schematic of a typical mass spectrometer leak detector

spectrometer leak detector is connected into the system, commonly in the backing line, whilst the ultrahigh vacuum pump of the system is either valved off or rendered inoperative. The system is pumped down by the backing pump and then the leak detection pumping unit (which in most designs is a diffusion pump/rotary pump combination) takes over. Some leak detector plants have a separate rotary pump incorporated for roughing out the system. When the pressure has dropped below say 10^{-3} Pa the spectrometer filament can be switched on, and the liquid nitrogen trap filled. Most units will function up to a maximum pressure of 10^{-2} Pa and down to 10^{-8} Pa. A similar dynamic range of leak rate can be measured. Suitable gauges such as a Pirani gauge for the higher pressures and an ion gauge for the low ultimate pressure measurements are included.

There is an alternative arrangement for mass spectrometer leak detecting, which exploits the dependence of the compression rate on the gas species in turbomolecular and diffusion pumps. In this arrangement, which Becker[6] has termed the 'counterflow' method, the system under test is connected to the foreline of the leak detector pump, whilst the spectrometer head is connected into the high vacuum side. The principle is illustrated in *Figure 8.4*. The pump discriminates by the compression ratio between the light test gas helium and the heavier gases H_2O, N_2 and O_2 which are the main gases present in a leaking system. As a result, more of the helium back diffuses through the pump than the other gases and the percentage of helium present in the high vacuum side will be many times greater than in the foreline side. Thus the pump acts as a filter. Although the performance of this arrangement is not necessarily better than the normal system, it has the advantages that it is simpler, does not require a liquid nitrogen trap and is less dependent on the pump speed. This

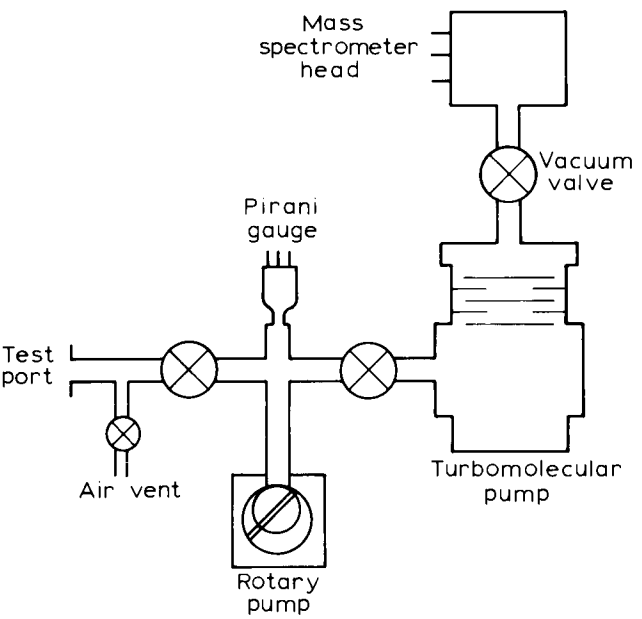


Figure 8.4 Arrangement of counter flow helium leak detector

type of system has been developed by Varian for their 936 family of leak detectors.

Whichever system is used, it should be pointed out that if the leak detector is to be connected to an ultrahigh vacuum system where freedom from oil vapour is essential, then the operator must ensure that oil from the rotary pump in the unit does not enter the vacuum system. If the unit does not have a liquid–nitrogen-cooled zeolite foreline trap, then it would be expedient to insert one between the unit and the system.

The sensitivity of mass spectrometer leak detectors depends to some extent on the pump speed as well as the design of the head and electronics. As will be discussed in the next section, standards have been proposed internationally for defining the performance of mass spectrometer leak detectors and most manufacturers, although not all, base their data on these recommendations. The sensitivity is usually defined at the leak rate giving 2% of full-scale deflection on the most sensitive range. Quoted values for air are around 5×10^{-12} Pa $m^3 s^{-1}$ at the normal pump speed which is usually about 10^{-2} $m^3 s^{-1}$. By throttling down the pump to reduce the effective pump speed, values of leak rate an order lower (5×10^{-13} Pa $m^3 s^{-1}$) can be detected. A commercial mass spectrometer leak detector is shown in *Figure 8.5*.

For very large vacuum systems, encasing the equipment with a helium gas bag or hood or indeed probing the whole of the outer surface with a small gas probe is often impractical. A method of overcoming this problem has been described by Solomon[7] in which the ratio of nitrogen-to-oxygen is measured with a mass spectrometer. If there is a leak, the ratio will be similar to that of the air surrounding the system, whereas the desorbed gases are likely to have a very different ratio. Thus the mass spectrometer can be used to enhance the pressure rise method mentioned in Section 8.1. Since large systems tend to have a number of isolation valves and pumps, these can be manipulated to pinpoint the area of the leak at which point a gas probe can be used.

8.7 Calibration of mass spectrometer leak detectors

Because of the parameters involved, the performance of a leak detector can be measured and also expressed in several ways. To compare equipment from different manufacturers, some agreement is required in presenting the data and with this end in view, an attempt has been made to draw up international standards on calibration procedure and data presentation. Although no final agreement has been reached, the draft proposal originating from the American Vacuum Society is generally acknowledged as a fair method and is adhered to by many leak detector manufacturers. The originating document, designated AVS Standard 2.1, defines the various terms and gives details of the calibration procedures to be followed.

For vacuum leak detection, the important term is the 'minimum detectable leak rate' (MDLR), defined as the smallest air leak rate that can be detected unambiguously by a given leak detector. Physically it depends on the effective pump speed, S_i, of the detector unit and the minimum partial pressure P_i of the search gas that can be detected, both referenced to the ion source.

$$\text{MDLR} = P_i \times S_i$$

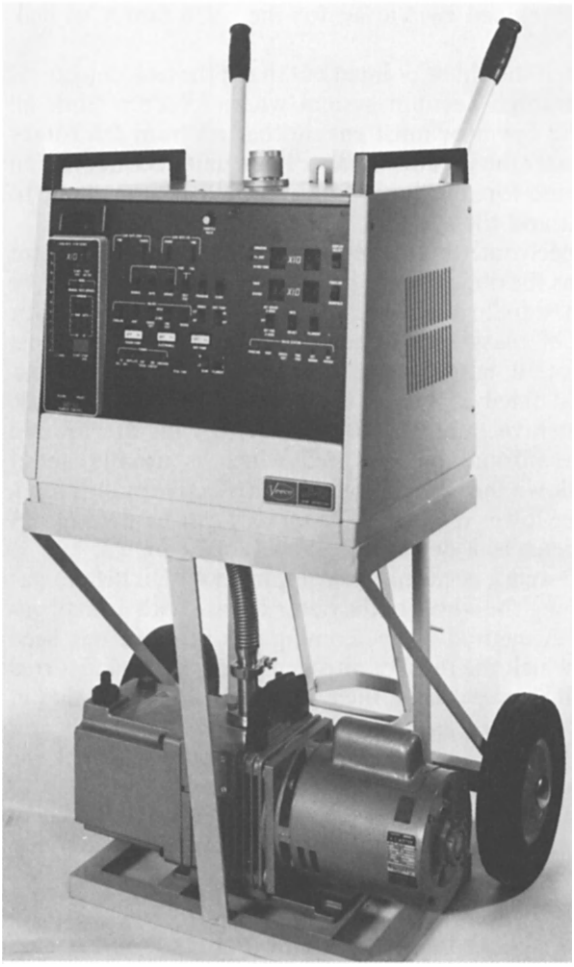


Figure 8.5 MS-20 leak detector with external roughing pump mounted on a transport trolley. (Courtesy Veeco Instruments Ltd.)

Sometimes MDLR is referred to as the sensitivity, but in the draft ISO proposal, sensitivity is reserved for defining the change in output of the detector amplifier divided by the change in input which causes the response. Thus the MDLR is calculated as the ratio of the minimum detectable signal to the sensitivity. The minimum detectable signal is taken as the sum of the noise and drift of the system with no test gas present, measured in scale divisions, or it is taken as 2% of full scale on the most sensitive range, whichever is the greater. The sensitivity is obtained by taking scale readings for known leak rates.

The proposal sets down the conditions and procedures for measuring the noise and drift and the sensitivity, and for converting the information into the MDLR for air, taking into account the background, the pump speed and the response time.

The equipment is given time to 'warm up' and to pump down before the

measurements are commenced and adjustments are made for optimum detection of helium. If the pump speed is adjustable, the selected speed should not be varied during the test. First the noise and drift are measured on a chart recorder over a period of 20 minutes. The maximum drift, in scale divisions over one minute, is designated the drift and the noise is taken as the largest deviation from the trace, multiplied by 2. Any abnormally large deviation which only occurs once during the 20 minutes can be considered as spurious and can be neglected.

For the sensitivity measurements, a source giving a known leak rate is required. Standard leaks are available commercially and usually consist either of a porous plug or of a membrane which can be connected to a helium supply. The membrane leaks usually have a helium bottle attached. The leak should be chosen to give a leak rate equivalent to about 50 times the minimum detector leak with the helium pressure at atmospheric. Normally the temperature will be quoted and a correction must be made if the measurements are made at a different temperature. To prevent any accumulation of helium in front of the leak an auxiliary pump system is required. This is illustrated in *Figure 8.6*.

In some leak detectors the auxiliary pumping system is incorporated in the unit. The auxiliary vacuum system is first tested for any background signal by closing tap V_1 or, if a valve is not fitted, replacing the calibrated leak with a plug. The auxiliary system is then pumped and connected to the detector by opening valve V_2 at the same time V_3 is closed. Any output reading is noted and represents a background correction. With V_2 closed, V_1 is opened to the leak, or, if no valve is fitted, the plug is changed for the leak, and the auxiliary system is pumped down. V_3 is then closed and V_2 opened. When the pressure in the leak detector has reached a steady value, the output signal is recorded. The leak isolation valve is then closed as fast as possible and a stopwatch

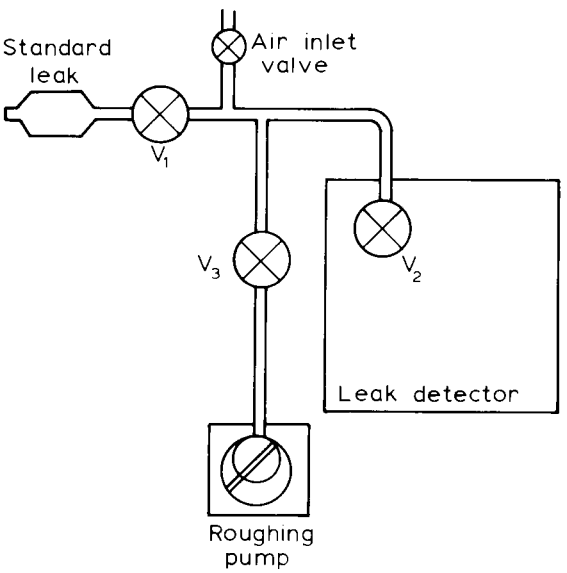


Figure 8.6 Test arrangement for measuring sensitivity of a mass spectrometer leak detector

started. The stopwatch is stopped when the reading has decreased to 37%. This time is taken as the response time. The sensitivity ϕ is then taken as

$$\phi = \frac{\text{Signal due to calibrated leak} - \text{Background}}{\text{Equivalent air leak rate of the calibrated leak}} \tag{8.26}$$

with a response time of T.

The MDLR with response time T is then

$$\text{MDLR}_{(T)} = \frac{\text{Minimum detectable signal}}{\phi} \tag{8.27}$$

For many leak detectors the noise and drift, i.e. the minimum detectable signal, are less than 2% of full scale on the most sensitive range and then the $\text{MDLR}_{(T)}$ is quoted in these terms, i.e. the leak rate which will give 2% of full scale on the most sensitive range.

The above gives an outline of the procedure. More detail in terms of the order in which steps should be taken and the time allowed for equilibrium are given in the AVS Standard 2.1 document. The document also details calibration of the minimum detectable concentration ratio of the test gas in air which is pertinent to larger leaks.

8.8 References

1. ROBINSON, N. W., *The Physical Principles of Ultrahigh Vacuum Systems and Equipment*, Chapman and Hall, Chapter 5, (1968)
2. POWELL, J. R. and McMULLAN, D., *Vacuum*, **28**, 287, (1978)
3. BLEARS, J. and LECK, J. H., Supplement to *J. Sci. Instrum.*, No. 1, 20, (1951)
4. KAMINSKY, M., *Atomic and Ionic Impact Phenomena on Metal Surfaces*, Springer-Verlag, Chapter 8, (1965)
5. WARMOLTZ, N., *Advances in Vac. Sci. Technol.*, (Proc. 1st Internat. Cong. on Vac. Tech., 1958), Pergamon Press, **1**, 257, (1960)
6. BECKER, W., *Vak. Technik.*, **17**, 203, (1968)
7. SOLOMON, G. M., *J. Vac. Sci. Technol.*, **A2**, 1157, (1984)

Appendix 1

Derivation of kinetic theory equations

A1.1 Pressure (Equation (1.2))

Because of the random motion of the molecules we can assume that there will be no preferred direction of motion for the molecules and the number of molecules per unit volume moving in a direction within a solid angle $d\omega$ will be given by

$$n\frac{d\omega}{4\pi}$$

where n is the total number of molecules per unit volume. If further we assume a velocity distribution function $f(v)$ such that $f(v)\,dv$ gives the fraction of the molecules with velocity between v and $v+dv$ in any direction, then we see that the number of molecules in a unit volume with a given velocity in a given direction will be

$$nf(v)\,dv\frac{d\omega}{4\pi}$$

Consider now a group of molecules with velocity v striking a small area of surface ds in time interval dt in the direction shown in *Figure A1.1*. These molecules will be contained at the beginning of dt in an oblique cylinder with base of area ds and slant length $v\,dt$. The volume of the cylinder will be $v\cos\theta\,dt\,ds$.

From the above, the number that will strike the surface in time dt, therefore, with this velocity and direction will be

$$nf(v)\,dv\frac{\sin\theta\,d\theta d\phi}{4\pi}v\cos\theta\,dt\,ds \tag{A.1}$$

The momentum of these molecules in the direction normal to the surface will be

$$mv\cos\theta\,nf(v)\,dv\frac{\sin\theta\,d\theta d\phi}{4\pi}v\cos\theta\,dt\,ds \tag{A.2}$$

Assuming elastic collisions the molecules will rebound at the same angle to

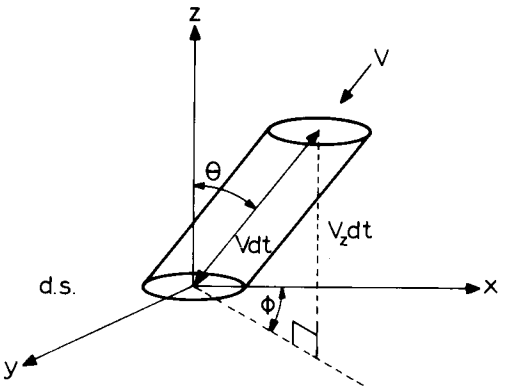


Figure A1.1 Molecules incident on a surface

the normal with the same velocity so that the change in momentum normal to the surface will be just twice the expression (A.2), i.e.

$$\frac{mn}{2\pi} v^2 f(v)\,\mathrm{d}v \cos^2\theta \sin\theta\,\mathrm{d}\theta\mathrm{d}\phi\mathrm{d}t\mathrm{d}s \tag{A.3}$$

Pressure is defined as the rate at which momentum is transferred to the surface per unit area and therefore the pressure exerted by the molecules in expression (A.1) is

$$\mathrm{d}P = \frac{mn}{2\pi} v^2 f(v)\,\mathrm{d}v \cos^2\theta \sin\theta\,\mathrm{d}\theta\mathrm{d}\phi \tag{A.4}$$

The total pressure is given by integrating over all directions and all velocities

$$P = \frac{mn}{2\pi} \int_0^\infty v^2 f(v)\,\mathrm{d}v \int_0^{2\pi} \mathrm{d}\phi \int_0^{\pi/2} \cos^2\theta \sin\theta\,\mathrm{d}\theta \tag{A.5}$$

which, since

$$\int_0^\infty v^2 f(v)\,\mathrm{d}v = v_{\mathrm{rms}}^2$$

by definition, gives

$$\begin{aligned} P &= nmv_{\mathrm{rms}}^2 \left[-\frac{\cos^3\theta}{3} \right]_0^{\pi/2} \\ &= \tfrac{1}{3} nmv_{\mathrm{rms}}^2 \end{aligned} \tag{A.6}$$

A1.2 Velocity distribution and mean velocity (Equation (1.7))

The Maxwell–Boltzmann distribution law can be applied to molecular velocities (see Chapter 5 of reference 2 in Chapter 1). This gives the velocity distribution in polar coordinates

$$f(\bar{v}) = \left(\frac{m}{2\pi kT}\right)^{3/2} \exp(-mv^2/2kT)v^2 \sin\theta \tag{A.7}$$

where $\bar{v}$ is the velocity in one direction.

The velocity function $f(v)$ for any random direction is given by integrating $f(\bar{v})$ over all directions, i.e.

$$f(v) = \int_0^{2\pi} \int_0^{\pi} f(\bar{v})\, \mathrm{d}\theta \mathrm{d}\phi$$

$$f(v) = \left(\frac{m}{2\pi kT}\right)^{3/2} 4\pi v^2 \exp\left(-\frac{mv^2}{2kT}\right) \tag{A.8}$$

The mean velocity is given by

$$v_m = \int_0^{\infty} v f(v)\, \mathrm{d}v$$

$$= 4\pi \left(\frac{m}{2\pi kT}\right)^{3/2} \int_0^{\infty} v^3 \exp\left(-\frac{mv^2}{2kT}\right) \mathrm{d}v$$

$$= \left(\frac{8kT}{\pi m}\right)^{1/2} \tag{A.9}$$

Comparing v_m with v_{rms} Equation (1.4)

$$v_m = \sqrt{\frac{8}{3\pi}}\, v_{rms} = 0.92\, v_{rms} \tag{A.10}$$

A1.3 Rate at which molecules are incident on a surface (Equation (1.8))

Expression (A.1) gives the number of molecules striking an element of surface in time $\mathrm{d}t$ with a velocity v in a given direction. The rate at which these molecules are incident on the surface therefore is

$$\frac{\mathrm{d}n}{\mathrm{d}t} = nf(v)\mathrm{d}v \frac{\sin\theta\, \mathrm{d}\theta \mathrm{d}\phi}{4\pi} v\cos\theta\, \mathrm{d}s \tag{A.11}$$

By integrating over all ranges of velocity in all directions the total number of molecules hitting the surface per second, v, can be deduced

$$v = \frac{n}{4\pi} \int_0^{\infty} v f(v)\, \mathrm{d}v \int_0^{\pi/2} \sin\theta\cos\theta\, \mathrm{d}\theta \int_0^{2\pi} \mathrm{d}\phi \tag{A.12}$$

Since $\int_0^{\infty} vf(v)\, \mathrm{d}v = v_m$

$$v = \frac{n}{2} v_m \left[\frac{\sin^2\theta}{2}\right]_0^{\pi/2}$$

$$= \frac{nv_m}{4} \tag{A.13}$$

A1.4 Mean free path (Equation (1.10))

The mean free path, λ, is defined as the average distance a molecule will go between successive collisions with other gas molecules.

We can calculate λ from the kinetic theory by considering a molecule of diameter d passing through the gas which is considered as stationary. The gas molecule will collide with another molecule when their centres come to within d of each other and we can therefore consider the moving molecule as having a radius of influence d, and consider the stationary molecules as points. The sphere of influence of the moving molecule traces out a cylinder of length v and cross section πd^2 in unit time if its velocity is v. The number of molecules it will collide with will then be $n\pi d^2 v$ where n is the molecular density. The problem is what value should be given to v to find the average number of collisions per unit time. Since the gas molecules are not in fact stationary the mean relative velocity v_{rm} is required. This can be calculated by averaging the relative velocity over the velocity distribution function of both molecules involved in each collision

$$v_{rm} = \iiint v_1 \, dv_1 \sin\theta_1 \, d\theta_1 d\phi_1 \iiint v_2 \, dv_2 \sin\theta_2 \, d\theta_2 d\phi_2 \, f(\bar{v}_1) f(\bar{v}_2) v_r$$

where the subscripts 1 and 2 refer to the two different molecules and v_r is the magnitude of the resultant velocity from $\bar{v}_1$ and $\bar{v}_2$.

The calculation is fairly lengthy and it is suffice to say here that the general result is given as

$$v_{rm} = (v_{1m}^2 + v_{2m}^2)^{1/2} \tag{A.14}$$

(See Reference 2 of Chapter 1.)

For a molecule travelling in its own gas $v_{1m} = v_{2m} = v_m$ and therefore

$$v_{rm} = \sqrt{2} v_m \tag{A.15}$$

The number of collisions therefore in unit time is

$$n\pi d^2 \sqrt{2} v_m$$

If the molecule is travelling at an average speed of v_m the mean free path will be given by

$$\lambda = \frac{v_m}{n\pi d^2 \sqrt{2} v_m}$$

$$\lambda = \frac{1}{n\pi d^2 \sqrt{2}} \tag{A.16}$$

Appendix 2

Vapour pressure curves of the elements

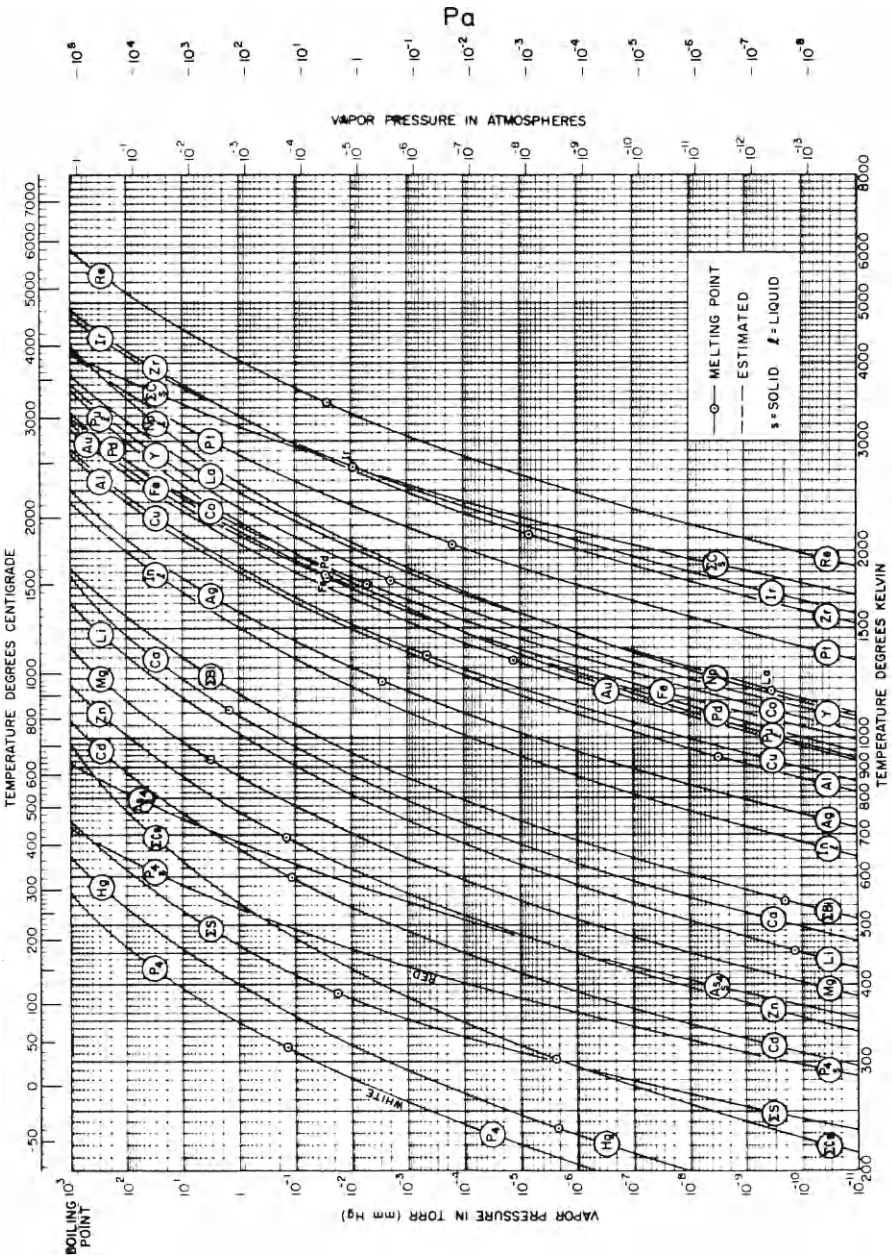


Figure A2.1 Vapour pressure curves of the elements (see Chapter 2, Honig[8])

Figure A2.2 Vapour pressure curves of the elements (see Chapter 2, Honig[8])

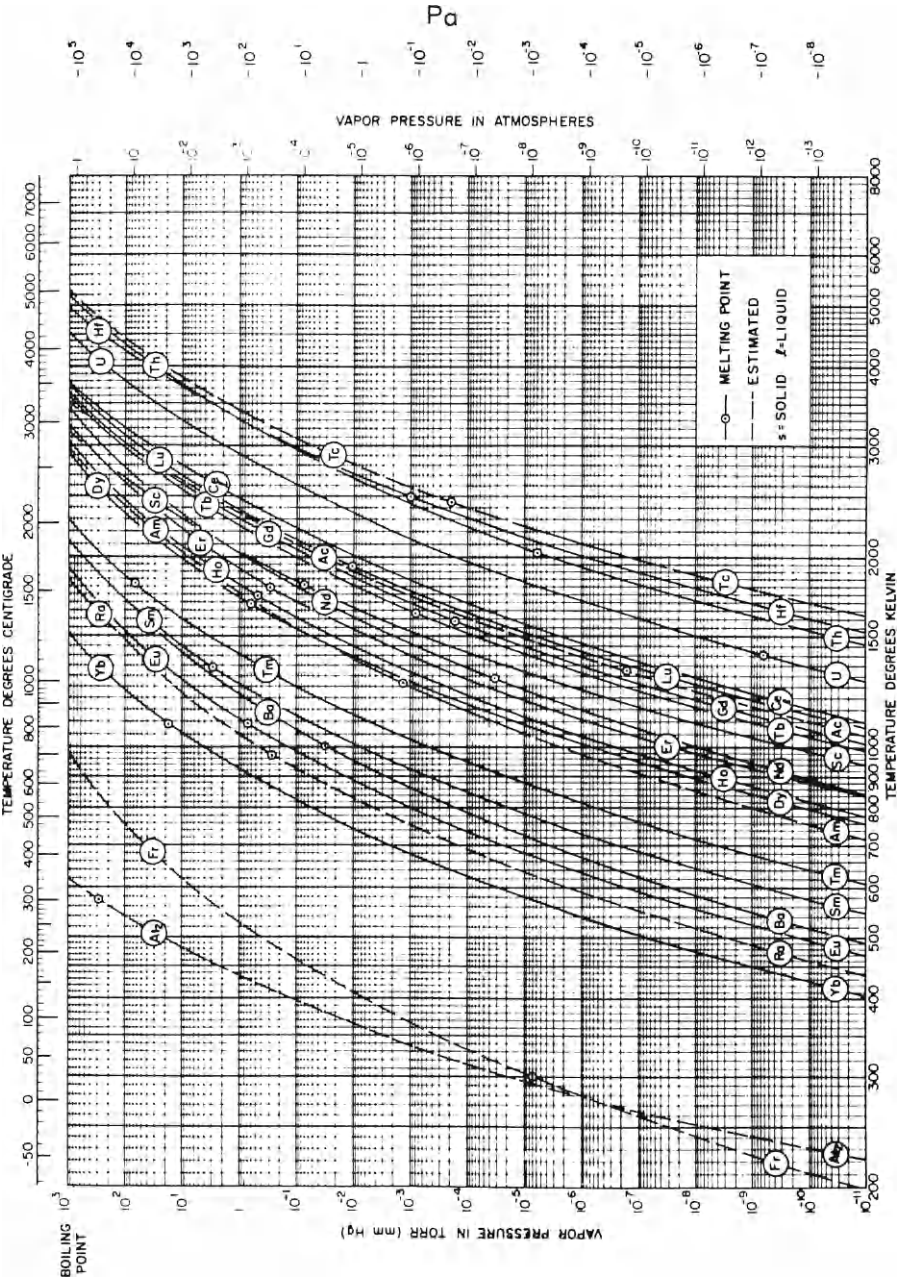


Figure A2.3 Vapour pressure curves of the elements (see Chapter 2, Honig[8])

Appendix 3

Conversion tables for vacuum quantities

Gas pressures

	Pa	Torr	mbar	*Standard atmospheres*
Pa or Nm^{-2}	1	7.5×10^{-3}	10^{-2}	9.87×10^{-6}
Torr or mm Hg	133.32	1	1.33	1.316×10^{-3}
mbar	10^2	0.75	1	9.87×10^{-4}
Standard atmospheres	1.013×10^5	760	1.013×10^3	1

Gas flow

	Pa m^3s^{-1}	Torr $l\,s^{-1}$	lusec	mbar $l\,s^{-1}$
Pa m^3s^{-1}	1	7.5	7.5×10^3	10
Torr $l\,s^{-1}$	0.133	1	10^3	1.33
lusec (μm Hg $l\,s^{-1}$)	1.33×10^{-4}	10^{-3}	1	1.33×10^{-3}
mbar $l\,s^{-1}$	0.10	0.75	7.5×10^2	1

Outgassing rates

	Pa $m^3s^{-1}m^{-2}$	Torr $l\,s^{-1}cm^{-2}$	lusec cm^{-2}	mbar $l\,s^{-1}\,cm^{-2}$
Pa $m^3s^{-1}m^{-2}$	1	7.5×10^{-4}	0.75	10^{-3}
Torr $l\,s^{-1}cm^{-2}$	1.33×10^3	1	10^3	1.33
lusec cm^{-2}	1.33	10^{-3}	1	1.33×10^{-3}
mbar $l\,s^{-1}cm^{-2}$	1.0×10^3	0.75	7.5×10^2	1

To convert from one unit to another, the units of the horizontal columns are multiplied by the table number to give the units heading the vertical columns. Thus to convert from torr $l\,s^{-1}cm^{-2}$ to Pa $m^3s^{-1}m^{-2}$ multiply by 1.33×10^3.

Index