CONTINUOUS PRESSURE CONTROL

THE ANALYSIS AND DESIGN
OF CONTINUOUS PRESSURE CONTROL SYSTEMS

BY

LOUIS BIALY,

B.Sc. (Eng.) (Rand).

A Thesis

Submitted to the Faculty of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Master of Engineering

McMaster University

(May) 1970

MASTER OF ENGINEERING (1970)
(Mechanical Design)

McMASTER UNIVERSITY
Hamilton, Ontario

TITLE: THE ANALYSIS AND DESIGN OF CONTINUOUS PRESSURE CONTROL SYSTEMS

AUTHOR: Louis Bialy, B.Sc. (Eng.) (Rand).

SUPERVISOR: Dr. M. C. de Malherbe

NUMBER OF PAGES: x, 278.

SCOPE AND CONTENTS: The continuous control of hydraulic pressure as a predetermined function of an independent variable is considered for the cases of low, intermediate and high pressure levels. A simple control system for the continuous regulation of low pressure is briefly discussed, and a means of extending its working range to intermediate pressure levels is described. Systems of this type are particularly suitable for the control of commercial diamond synthesis presses, but may also be utilised for other processes.

A more sophisticated approach to problems of pressure generation and control is discussed in detail, with particular reference to a high pressure isostatic press. Digital computer methods of optimising the parameters of the system and simulating its response are developed.

ACKNOWLEDGEMENTS

The value of this study has been mainly enhanced by the efforts of Dr. M. C. de Malherbe, for whose continual guidance and encouragement, the author is grateful.

The author is indebted to Dr. N. C. Pope, Assistant General Manager, Hard Metals Ltd., and Mr. K. O'Donovan, Manager, Development Division, Industrial Distributors (1946) Ltd., for their personal interest, and sincerely appreciates their efforts in making this study possible.

The financial support provided by Industrial Distributors (1946) Ltd., Republic of South Africa, is gratefully acknowledged.

TABLE OF CONTENTS

TABLE OF CONTENTS cont.

TABLE OF CONTENTS cont.

TABLE OF CONTENTS cont.

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS cont.

LIST OF ILLUSTRATIONS cont.

1. INTRODUCTION

Many scientific and industrial processes require the application and accurate regulation of high and very high hydraulic pressures. This study is concerned with the analysis and design of systems suitable for the continuous control of pressure as a function of time or some other independent variable.

Although the field of High Pressure Engineering is well established as a discipline, there is a lack of agreement in the literature as to the pressure ranges denoted by the terms 'high pressure', 'very high pressure' and 'ultra high pressure'. For the purpose of this study the terms defined in Table 1 are adopted.

TABLE 1 . DEFINITION OF PRESSURE TERMS

TERM		PRESSURE RANGE
Low pressure	L.P.	Below 3,000 p.s.i.
Intermediate pressure	I.P.	3,000 p.s.i. to 10,000 p.s.i.
High pressure	H.P.	10,000 p.s.i. to 100,000 p.s.i.
Very high pressure	V.H.P.	100,000 p.s.i. to 400,000 p.s.i.
Ultra high pressure	U.H.P.	Above 400,000 p.s.i.

Each of the pressure ranges defined in Table 1 imposes different pressure generation, control, and other problems. A different approach is then generally required in each case.

Continuous pressure control in the L.P. range may conveniently be effected using standard components only, [see Chapter 4] while the same basic circuit with one non-standard component included will suffice for control in the I.P. range [see Chapter 5]. For the H.P. range, intensification of hydraulic pressure is necessary and a new approach to the control problem is required if reliability and efficiency are to be achieved [see Chapter 6]. In the V.H.P. range special fluids are required to prevent freezing while more than one stage of pressure intensification will generally be necessary. The problem of continuous pressure control in the V.H.P. range is briefly examined in Chapter 7. The problem of fluids freezing precludes their use in the U.H.P. range. Solid media must be used for load transmission and final pressure amplification, but control may still be effected by regulating an intermediate, lower pressure, hydraulic stage [see Chapter 2].

Several applications of continuous pressure control are considered and some problems directly and indirectly associated with the synthetic diamond industry are identified in Chapter 2. The bulk of this study is devoted to the solution of these problems. Most of the previous work done on pressure control has been devoted to devices for steady pressure regulation. Some of these devices are reviewed in Chapter 3.

2. SOME APPLICATIONS OF CONTINUOUS PRESSURE CONTROL

2.1 Outline

Several applications of continuous pressure control are briefly considered. Since this study was initiated by the demands of the Synthetic-Diamond industry the requirements of this field are more carefully examined and some problems are identified.

2.2 Hydrostatic Forming

Known variously as 'Fluid Forming', 'Hydrostatic Pressing' and 'Hydrostatic Forming' the process makes use of hydrostatic fluid pressure to perform manufacturing operations such as drawing, deep drawing, and hydroblock or cavity die forming. These manufacturing processes are all associated with the sheet metal working industry.

An hydrostatic press capable of operating in the I.P. and H.P. ranges to pressures of 16,500 p.s.i. is described in reference [1]. The press consists of two vertically opposed L.P. Rams (see Fig. 1) in an Asea Quintus 3300 ton press frame. The "fluid-forming" dome is attached to the upper ram. The dome consists of a rubber bag enclosed on the sides and top by a steel cylinder. The exposed surface of the bag is the working portion of the press and comes in direct contact with the material being formed.

After insertion of the die and material to be formed the dome is clamped against the press table using the upper L.P. ram. The rubber bag is thus contained in all directions and this prevents rupture of the bag when it is pressurised. The bag is connected directly to the I.P. - H.P. source via a floating supply line.

The press may be used as a conventional downstroking press with the lower ram used only as an ejector if required. It may also be used as a double acting press with solid tools as in a conventional draw. The major advantages of this type of press become apparent when it is used with the "fluid-forming" dome either single acting for cavity die forming or double acting for deep drawing.

For deep drawing operations a pressure - displacement cycle is required. In reference [1] the dome pressure as a function of the lower ram stroke is carefully regulated by a preset programme. The programmer in this case is a paper cam which is operated by the lower ram and followed by a photo-electric follower which actuates the pressure control system. The actual pressure control system is protected by an Asea patent and is not discussed in the publication.

2.3 Organic Chemical Synthesis Under Isostatic Pressure

The terms "Isostatic pressure" and "Hydrostatic pressure" are often used interchangeably in the literature. Some authors do however draw a distinction. For the purpose of this study the following definitions are adopted:-

(i) A material is considered to be under isostatic pressure when it, (or the flexible container in which it is enclosed) is completely surrounded by a pressurised fluid.

(ii) A material is considered to be under <u>hydrostatic pressure</u> when it, (or the flexible container in which it is enclosed) is partially surrounded by a pressurised fluid.

These definitions, though somewhat arbitrary, are in agreement with definitions adopted by other authors in the high pressure field. (e.g. see reference [3].)

Industrial chemical synthesis under isostatic pressure is limited at present to pressures under 50,000 p.s.i., however there is evidence to suggest that such processes could be carried out at pressures as high as 300,000 p.s.i. [4].

Basic research in high pressure chemistry is currently being executed in the H.P. and V.H.P. ranges. Research in the V.H.P. range is envisaged for the near future [2].

Thomas et al [2] describes a piston-cylinder apparatus suitable for generating pressures of up to 360,000 p.s.i. for laboratory work on organic chemical synthesis. Basically the apparatus (see Fig. 2) consists of a V.H.P. cylinder and piston mounted between the plattens of a L.P. press. The working portion of the cylinder constitutes a controlled volume of up to about 3.5 cubic in.

The reactants are encapsulated in a flexible metallic container resembling a toothpaste tube which may be made from tin, lead or annealed gold tubing as required. The container is inserted in the V.H.P. chamber and after filling with oil the region is pressurised by forcing the piston into the cylinder using the L.P. ram of the press. The V.H.P. cylinder pressure may be controlled by regulating the pressure in the L.P. cylinder.

External heating of the V.H.P. cylinder may be applied to ensure that the oil does not freeze at the pressure level to be used.

While adequate for small scale laboratory tests, this apparatus would not be suitable for large scale industrial chemical synthesis. The reasons for this are fully discussed in Chapter 6.

It is entirely possible that in the foreseeable future organic chemical synthesis processes requiring continuous pressure - temperature - time cycles will be utilised on a commercial scale in the H.P. and V.H.P. ranges.

2.4 Hydrostatic Extrusion

The hydrostatic extrusion process is illustrated in Fig. 3. A billet of material to be extruded is placed in the cavity of the extrusion cylinder, and forced through the orifice of a suitable die at one end by the action of hydrostatic pressure. The pressure is generated in a suitable fluid which is enclosed in the chamber, by a plunger which is driven by a lower pressure ram. Pressures employed are usually in the V.H.P. range, and are in fact limited only by the strength of container required and by the problem of fluid freezing at excessive pressures.

The major advantage of hydrostatic extrusion is the lack of friction between the billet and its containing vessel [5].

From a control point of view, one of the major problems is the lack of control over the rate of extrusion. This problem has been extensively discussed throughout the literature [5, 6, 7, 8, 9]. When the plunger is moved into the vessel at a constant velocity, a uniform rate of extrusion would be expected. A 'stick-slip' movement of the extruded product has however been observed in the majority of cases. This undesirable phenomenon is accompanied by a fluctuating, and often oscillatory pressure-time characteristic [6]. Bridgmen [9] concluded that the 'stick-slip' phenomenon is an inherent limitation of the process. Several authors have however investigated the 'slip-stick' characteristic further and have in many cases devised means of limiting it.

It is apparent that the 'slip-stick' phenomenon is not fully understood but it is generally agreed that the major cause is the unusual billet-die friction characteristics [6, 7, 10]. When the billet first contacts the die, the interface pressure exceeds the fluid pressure due to the effect of the 'unsupported area' behind the orifice of the die. This high interface pressure prevents the penetration of lubricant [7]. When the hydrostatic pressure surrounding the billet builds up sufficiently, extrusion is initiated and relative slip between the billet and die surfaces takes place, and a film of lubricant is drawn in, thus reducing the apparent coefficient of friction. This implies that a lower pressure would be sufficient to maintain extrusion. The excess pressure accelerates the material through the die. Due to stored energy in the compressed fluid, the pressure falls off relatively slowly. Crawley, Pennel and Saunders [7] suggest that the apparent coefficient of friction is a decreasing function of velocity, thus as the velocity of extrusion continues to increase, the decreasing friction accelerates the process. Thus it is possible that the process is still accelerating even after the pressure has fallen below the mean extrusion pressure. (The pressure drop is caused by the fact that material is leaving the vessel at a greater rate than fluid is being displaced by the plunger.) When the pressure approaches the minimum extrusion pressure the rate of extrusion begins to decrease and thus the friction increases which quickly brings the process to a halt. The cycle is then repeated and the net result is an unstable, oscillatory motion. It should be emphasised that in a typical process the transients are very rapid, of the order 10 ms. for a pressure cycle [7]. The plunger which necessarily is a high inertia device simply cannot cope with transients of such short duration.

Several solutions to the problem have been proposed. Lowe et al [6] show that under certain conditions, the control of extrusion speed may be improved by choosing the correct die entry angle, minimising the volume of compressed fluid, roughening the billet surface to improve lubrication and modifying the billet nose to initiate extrusion at an early stage in the pressure cycle. The authors also propose an alternate solution to the problem. This involves the fitting of a damper to the end of the billet. The damper is merely a short length of round bar recessed at one end to fit snugly over the end of the billet (see Fig. 4). The recess is sealed by means of an 'O' ring and the damper held in position by the 'unsupported area' principle. The damper has several fine axial holes drilled through it and is a free running fit in the cylindrical pressure vessel.

During stable conditions the damper moves through the fluid offering little resistance to billet movement. However, a sudden acceleration of the billet causes an increase in pressure ahead of and a decrease in pressure behind the damper which can generate a significant braking effect. This device proved successful under a variety of conditions, particularly for conical entry dies of up to 60° included angle. It has been unsuccessful in suppressing instability during extrusion through dies with greater angles of entry.

Slater and Green [5] found that 'slip-stick' motion is damped down and sometimes obviated altogether by using augmented hydrostatic extrusion. In this process the vessel is pressurised to a value somewhat less than the extrusion pressure, and extrusion is effected using a separate ram in very much the same manner as in the conventional extrusion process. The process thus requires an external source of V.H.P. Augmented extrusion is necessarily a more expensive process than simple hydrostatic extrusion, however should it prove economically feasible, then the possibility of using external control of pressure should be considered. A continuous pressure control system similar to that described in Chapter 6 could provide a final solution to the slip-stick problem, and could be used in the case of simple hydrostatic extrusion as well. The system would function as a velocity control system, the difference between the required and actual extrusion speeds being used to regulate the hydrostatic pressure. The plunger velocity could be simultaneously controlled by an auxilliary circuit.

The design of a velocity control system would require a complete dynamic analysis of the hydrostatic extrusion process. Crawley et al [7] develop a rather sophisticated mathematical model of the process, which may be reduced to a system of eighteen first order differential equations. The simultaneous solution of these equations may be executed on a digital computer using a method similar to that presented in Chapter 6.

2.5 Autofrettaging

The benefits that the autofrettaging of thick walled pressure vessels can provide are well established [11, 12, 13, 14]. When a cylinder is subjected to an internal pressure, the absolute magnitude of the elastic stress distribution in the wall of the cylinder is a decreasing function of radius. This is true for both radial and circumferential principal directions. As a consequence of this the material strength of the outer region of the cylinder is very poorly utilised.

If the cylinder is initially overstrained by the application of an internal pressure of sufficient magnitude, the bore of the cylinder will suffer a plastic deformation. When the internal pressure is relaxed, the tension in the elastic portion of the cylinder will induce a residual compressive stress in the portion of the cylinder which was plastically deformed. In subsequent applications of internal pressure a more favourable stress distribution will result since the residual compressive stress must be overcome before the circumferential stress can become tensile.

Ideally, the material should be at the point of reverse yeilding, i.e. yielding in compression at the end of the process [14]. This imposes a limit on the overstraining pressure which may be used; furthermore the collapse pressure of the cylinder [14] must not be exceeded.

Various methods of predicting the maximum autofrettaging pressure to be used have been put forward [12, 13, 14]. The pressures required are usually in the H.P. and V.H.P. ranges. It has been reported [2] that the use of these methods to predict the pressure at which plastic deformation first occurs, and the pressure at which the process should be stopped does not always give reliable results. This means that in many cases the autofrettaging process must be empirically performed. In practice it is often done in stages, the internal bore diameter being used as measure of the plastic deformation. This approach is expensive and does not lend itself to mass production.

The efficiency of the process may possibly be improved by utilising an automatic continuous flow - pressure control system. If for example the rate of flow of fluid into the vessel to be autofrettaged is kept constant by an accurate feedback control system, the time rate of pressure rise will be constant as long as Hooke's law is obeyed. When plastic yielding of the cylinder commences, the slope of the pressure - time curve will begin to decrease and will continue to do so as long as plastic yielding persists. The slope of the pressure - time curve may thus be used as a measure of the fraction of the cylinder volume which has been plastically deformed. When the slope reaches some lower limit which has been empirically determined, the pressure may be automatically released. The pressure and rate of pressure rise would have to be carefully monitored by the control system. In an optimum system of this nature, the flow rate need not be kept constant, but could be regulated to ensure maximum hydraulic power transfer to the cylinder. (The hydraulic power transfer would of course be limited by available hydraulic power, and subject to a satisfactory transient pressure response.) The cut off - pressure slope would have to be adjusted accordingly, and a continuous pressure-flow control system would be necessary.

2.6 High Pressure Cutting

Although generally applied to softer materials such as plastics, wood and leather, it is believed that high pressure cutting will be utilised on a commercial scale on harder materials including ferrous metals and ceramics in the future.

The process consists of directing a very narrow jet of H.P. fluid (usually water) on the material to be cut. The fluid pressure upstream of the nozzle governs the depth of cut while the flow rate determines the maximum allowable feed rate. A multitude of variations on the basic process are possible, e.g. embossing, lithographing, engraving, surface treatment of metals, drilling, etc. The major advantage of the process at present is considered to be its efficiency over methods such as sawing and shearing in certain applications.

Applications of the process which require control over the depth of jet penetration may utilise a continuous pressure regulation system. In cases where the actual depth of penetration lends itself to measurement, the error between the instantaneous required and actual depths may be used to correct the pressure upstream of the nozzle and hence the velocity of the jet.

2.7 Synthesis of Diamond

Diamond boart is synthesised on a commercial scale by subjecting graphite, in the presence of a catalyst to a special pressure - temperature - time cycle. For the process to be commercially feasible V.H.P.s of the order 65 Kilo-bars (Kb)* (950,000 p.s.i.) and temperatures of the order

---

* 1 bar = 14.5 p.s.i., 1 Kb = 1,000 bars.

1700°C are required. Pressures of this magnitude cannot be hydraulically generated due to the fact that all fluids, (even at elevated temperatures) freeze at much lower pressures than this. Commercial diamond synthesis processes thus make use of solid materials for the final stage of pressure amplification and transmission [15, 16].

The commercial diamond synthesis process is carried out in an hydraulic press, which may have anything from a single ram to eight rams. The graphite and necessary catalysts are enclosed within a capsule, the shape of which is determined by the number of rams to be used. The ram load is transmitted to the capsule by means of "pressure magnifying anvils", which in the ideal case amplify the cylinder pressure by the ratio of cylinder area to the 'nose area' of the anvil.

Among the many possible press configurations, the most successful are the 'cubic-type' and the 'vertically opposed anvil' type. The cubic press consists of six rams each fitted with a suitably shaped anvil. The anvils act on a capsule which is in the form of a cube. Some of the capsule material extrudes into the gaps between the anvils, under the action of the high pressure, thus sealing the volume and containing the pressure. One of the major difficulties with presses of this type is the synchronisation of the rams [17]. Even when accurate feedback control systems are used, slight differences in the damping of each ram (e.g. due to differences in friction) can lead to transient out of balance loading on the capsule. The stability of the extruded seal is very limited and is easily upset thus causing the pressure to "blow-out". The "blow-out" is often accompanied by damage to the anvils and peripheral equipment. Recent developments [17], at H.P.M.L., in the Republic of South Africa, indicate that this problem has been largely overcome. However, due to its complexity the cubic type press still represents a sizeable capital outlay which must be weighed against the advantages it offers over other configurations. These will be considered subsequently.

The press configuration that has proved to be economically most successful makes use of the 'Bridgman Opposed Anvil' system, sometimes referred to as the 'Belt System'. It was in a configuration of this type that diamond was first synthesised in Sweden in 1952, and independently in the U.S.A. in 1955 and South Africa in 1957 [15]. The 'opposed anvil' apparatus is discussed by Bridgman [18] while the original 'belt' is described by Hall [19]. An updated version of the 'belt' system is illustrated in Fig. 5.

The high pressure generated in the capsule by the anvils is contained by the die or belt. The anvils and dies are supported where possible by concentric binding rings for reasons which are subsequently discussed. The pressurised volume is sealed as in the case of the cubic by the 'gasket' formed by the extrusion of the 'wings' (Fig. 6) of the capsule material into the narrow gap between anvil and die. The seal is far more stable in this case than in the cubic since there are fewer surfaces, and anvil motion is in one direction only.

The high temperature required for synthesis is achieved by electric resistance-heating of the capsule. The graphite and catalyst materials are good conductors of electricity, hence a low voltage (of the order 10 volts), heavy direct current (1,000 to 5,000 amps.) is applied directly across the anvils, which must necessarily be electrically insulated from the press.

The capsule is required to fulfill the following functions:-

(i) To contain the graphite and catalysts.

(ii) To provide a stable pressure seal between anvil and die.

(iii) To electrically insulate the die from the graphite and the anvils.

(iv) To provide as much thermal insulation as possible to minimise temperature gradients within the contents of the capsule.

A typical capsule is illustrated in Fig. 6. One of the most successful materials for capsule manufacture is an unusual plutonic rock known as pyrophyllite or 'wonder-stone'. It is easily machined, and displays quasi-plastic properties under the action of high pressures. This material may be heat-treated to improve its extrusion properties. It is also a good thermal and electrical insulator.

It is necessary to cool the contiguous cylinders which support the anvils and die, in order to minimise thermal stresses and to permit better temperature control within the capsule. Cooling is achieved by passing water through jackets on all exposed surfaces of the binding rings.

The die is lined up with the anvils using an alignment sleeve, and the assembly of anvils, die, supporting rings, water jackets and alignment sleeve is referred to as a 'die-set'.

The anvils are required to withstand compressive stresses of the order 65 Kb. They are thus made from a material with a very high compressive strength, such as a tungsten-carbide alloy containing between 5% and 15% cobalt. Pure compressive tests on tungsten carbide specimens indicate that apparent ultimate compressive strengths in excess of 60 Kb are unusual [20].

Higher apparent ultimate compressive strengths are however obtainable if the so called "massive support principle", originally proposed by Bridgman [18] is used. As an example of this principle Bridgman cites the case of two heavily supported knife-edges being pressed against each other. Much higher compressive stresses at the contact areas are possible than the material normally can withstand due to the support offered the small compressed regions by the massive amounts of surrounding material.

The anvils are designed with a very much larger 'base' than 'nose' area and thus the "massive support principle" applies [21]. The anvils are reinforced by a series of concentric steel binding rings which are shrunk onto the cylindrical base of the anvil.

In the actual diamond synthesis process the anvils are also subjected to severe thermal stresses due to high temperature gradients.

A lot of research has been devoted to optimising the physical dimensions and shape of anvils in order to maximise their strength. As a result of this work at the High Pressure Materials Laboratory and other institutions associated with the synthetic diamond industry in South Africa the design of the Ultra-high pressure anvils has been optimised. Efforts to improve the properties of the material in order to prolong anvil life are still continuing. This aspect is further discussed subsequently.

Much of what has been related regarding anvils applies to dies as well. The dies are also made from tungsten carbide to resist the high compressive stresses. The design of the die and support rings appears however, to be a more exacting problem than that represented by the anvil. Since tungsten carbide is very weak in tension, the concentric steel binding rings have to lend sufficient support to ensure that the tensile stresses in the die never exceed the elastic limit. In an optimum design of this nature, the inner tungsten carbide liner would be precompressed to such an extent that the material at the bore would be at the point of yielding in compression. When subsequently subjected to V.H.P., the die would undergo a complete reversal in circumferential bore stress to almost reach the elastic limit in tension.

Methods of optimising the geometry of the contiguous cylinders, and the interferences to be used have been proposed by many workers in this field [14, 24, 25, 26, 27].

Despite the use of optimum designs for the 'belt', it still remains the weak link in the diamond synthesis process. Although the process is still economically feasible, the toll in broken dies (and to a lesser extent broken anvils) does not go unfelt.

One of the major advantages of the 'cubic' type of press is that the belt is replaced by four active anvils which do not suffer from such severe tensile stress problems.

Alternative designs of the 'belt' apparatus which would include 'end loading', i.e. support in the axial direction as well, are being examined. The principle behind this approach is that if conditions approaching 'all round' hydrostatic pressure on the inner liner of the 'belt' can be achieved, then this liner can only suffer elastic volume contractions, and its strength in compression or tension will be of minor importance [23].

As in the case of the anvil, methods of improving the material properties of the tungsten carbide die are being investigated. Such investigation is being carried out on a broad front from basic metallurgical research in the properties of tungsten carbide alloys, to a continuous review of the manufacturing processes used in the production of anvils and dies. Some of the aspects of the manufacturing process relevant to this study are discussed in section 2.8.

One of the most important requirements for economical manufacture of synthetic diamonds is that the encapsulated material undergoes the empirically determined pressure - temperature - time cycle. Significant deviations from the ideal cycle can seriously reduce the 'percentage yield' of the process. (The 'percentage yield' is the percentage by weight of graphite in the capsule converted to diamond.) The pressure and temperature cycles are synchronised and controlled by a single programmer. Consideration of the temperature control system is beyond the scope of this study.

It is desirable to use as large a capsule as possible to maximise the yield per run of the process. Larger capsules also lead to higher 'percentage yields' since less severe temperature gradients are present. There is therefore a tendency towards the use of presses with ever increasing load capacities so that larger capsules may be accommodated.

Presses in current operation have capacities between 3,000 and 10,000 tons.* For presses of this nature the size of the working cylinder and hence the maximum cylinder pressure is governed by economic considerations. Modern press-frame design and the manufacturing costs associated with cylinders of large bore favour compact cylinders with the result that maximum working pressures in the I.P. and H.P. ranges have to be used. As an example, a 10,000 ton press would require a cylinder with a 38 inch bore if a pressure of 20,000 p.s.i., is used. If the maximum pressure were restricted to 3,000 p.s.i., the cylinder would have to have an eight foot bore.

Presses currently in use generally utilise maximum cylinder pressures from about 6,000 p.s.i. to about 20,000 p.s.i., although there is still some demand for lower pressure units.

Due to practical difficulties in measuring the pressure within the capsule during production runs, the cylinder pressure is generally used as a measure of the capsule pressure, and effects of seal friction and other disturbances are taken into account in the design of the pressure - time programme. This implies that there will be differences in the optimum cycle used on each press and die set.

The pressure - time cycle employed is generally complex, but of a continuous nature, with maximum pressure rise and fall rates of up to 150 p.s.i./sec. The maximum allowable error in cylinder pressure at any instant in time is usually specified as a percentage of the pressure range used. (Usually 0.1% to 0.5%.)

---

* 1 ton = 2240 lb.

A system suitable for the continuous control of pressure in the L.P. range is analysed in Chapter 4, while a means of extending the maximum working pressure of the system into the I.P. range is described in Chapter 5.

2.8 Isostatic Pressing

Isostatic pressing is a technique which is utilised in the powder metallurgy and ceramic industries and is beginning to find applications in other fields such as the chemical and plastic industries [3]. It is used to compact and densify powders into coherent masses.

The process involves enclosing the powder in a flexible container and inserting this loaded container into a suitable pressure vessel (Fig. 7).

The chamber of the vessel is subsequently subjected to hydraulic pressure in the I.P. range, and on decompression the flexible container recovers its original shape and the compact may be removed.

This technique lends itself readily to the production of compacts in special shapes, e.g. anvils and dies for synthetic diamond manufacture, and avoids the lack of uniformity of density and cracking due to frictional losses between powder and die wall associated with conventional pressing methods.

The flexible container or 'bag' is usually made from plastic or rubber. Rolfe [3] recommends that the bag be made from neoprene or nitrile rubber if oil is the working fluid, and butyl rubber if water is to be the working fluid. Some manufacturers recommend P.V.C. as an all purpose container material [28]. The shape of the bag governs the shape of the compact, hence the necessity for accurate bag manufacture. The bag should be sufficiently rigid to maintain its shape during loading.

After loading, the powder is covered with a piece of filter paper before the lid is placed on the container. The container is then evacuated to a pressure of about 1 torr* and sealed. The lid of the bag may be so designed that the hydraulic pressure assists in maintaining the seal.

In the manufacture of tungsten carbide dies and anvils, the constituents are thoroughly mixed using a milling process and the powder isostatically pressed as outlined above. The compacts are then sintered under carefully controlled conditions in hydrogen furnaces. It is believed that the subjection of the material to a continuous pressure - time cycle during the isostatic compaction process can significantly improve the properties of the finished product and lead to longer mean component lives.

The design of a continuous pressure control system for isostatic presses is considered in Chapter 6.

---

* 1 torr = 1 mm Hg.

3. STEADY PRESSURE REGULATION

3.1 Outline

A review of the literature reveals that most previous work in the field of hydraulic pressure control is devoted to steady pressure regulation devices. These devices may be broadly classified as pressure limiting valves, minimum pressure valves and pressure reducing valves. Some of these devices are briefly discussed.

3.2 Pressure Limiting Devices

The purpose of pressure limiting devices is to prevent overpressures in hydraulic circuits. Such devices are particularly necessary when fixed-displacement pumps are used [29]. Many pressure limiting devices have been devised to provide general and specialised protection to hydraulic circuits and components [30, 31]. The most widely used devices are relief valves, safety valves and pressure fuses.

Pressure fuses are simply metallic diaphragms arranged at crucial points in an hydraulic circuit. These diaphragms are designed to rupture if a given pressure is exceeded. A pressure fuse is the hydraulic equivalent of fuse-wire in an electric circuit.

Safety valves are the hydraulic equivalents of circuit breakers in electrical power distribution. When a preset pressure level is exceeded, the safety-valve is triggered and directs all flow through the valve to the reservoir. A safety valve must in general be manually reset as it does not reset itself when the pressure returns to normal.

Safety valves are similar in construction to relief valves except that they have some form of 'hold device', which is fail-safe in operation. Further consideration of safety valves is not warranted.

Relief valves are by far the most important and most widely used of the pressure limiting devices.

Many different types of relief valves are available [29, 30, 31] but the principle operation of most forms is similar. A typical single-stage pressure relief valve is schematically illustrated in Fig. 8.

The pressure to be controlled is the supply pressure P_s, which is sensed on the spool end area and compared with the spring force setting. The force differential actuates the spool valve which controls the flow bypassed to the reservoir thus tending to maintain the supply pressure at the set value. The desired supply pressure may be attained by adjusting the spring force setting.

The restriction in the sensing line is necessary in order to stabilise the valve [32, 33]. In a dynamic analysis of the relief valve, Merritt [33] shows that if the restriction is absent, the open-loop frequency response is dominated by the spool valve mass-spring dynamics. In practice the spool is very lightly damped by viscous friction and instability results. The effect of the restriction is to provide rate feedback, which compensates for the low damping ratio and stabilises the system. The restrictor also introduces a low frequency lag caused by the interaction of its hydraulic resistance and the equivalent capacitance of the sensing chamber volume. This affects the transient response of the relief valve to sudden changes of flow. Thus large supply pressure overshoots may occur while the pressure P_c is building up.

Two stage relief valves are often used in hydraulic circuits. These devices operate on the same principle as the type described in the foregoing but have an additional pilot stage. They are generally more accurate than single stage types, but have slower responses and are less reliable as they are more susceptible to clogging [33].

3.3 Minimum Pressure Valves

While pressure limiting devices are primarily intended for use in a protective capacity, minimum pressure valves are active logic components in hydraulic circuits. Many different types of minimum pressure valves are available [29, 30], the most widely used being sequence valves and unloading valves.

Sequence valves are used to direct fluid to more than one branch of an hydraulic circuit in sequence. Flow is directed only to the primary circuit until a preset primary pressure is reached. At that stage the valve commences delivery to a secondary circuit. The pressure in the secondary circuit may be any required fraction of that in the primary, but may never exceed it. The primary pressure is however independent of the secondary pressure. A simple sequence valve is schematically illustrated in Fig. 9. If the secondary pressure is very low, a large secondary flow will take place, causing a drop in primary pressure, which tends to close the valve, and reduce the secondary flow. This can lead to instability as in the case of the relief valve and hence the necessity for the restrictor in the pilot pressure line.

Unloading valves are usually used to direct primary flow to the reservoir once a desired pressure in some other part of the circuit has been reached. A good example is the unloading of a pump as soon as a cylinder or accumulator has reached its full pressure. In construction an unloading valve is almost identical to a sequence valve. However, the pilot pressure port is connected to a remote source rather than to the supply pressure, while the secondary port is connected to the reservoir. The restriction is unnecessary in this case since the valve unloads for as long as the pilot pressure exceeds the pressure setting of the valve. Unloading valves therefore respond far more rapidly than any of the other types.

Both unloading and sequence valves are commercially available in two-stage versions, which are more accurate and smoother in performance than single stage types.

3.4 Pressure Reducing Valves

Fluid power applications often require several different pressures from a single pressure source. The utilisation of pressure reducing valves provides a convenient means of achieving this end. Pressure reducing valves are designed to supply lower pressure fluid to a branch circuit from a higher pressure line.

A single stage pressure-reducing valve is schematically illustrated in Fig. 10. It is similar in construction and operation to a relief valve, the major difference being that the reduced pressure, P_r rather than the supply is sensed and compared with the reference spring force setting.

Any deviation from the reference pressure will cause a change in the valve opening and hence the flow through the valve. The flow varies in such a way as to minimise the deviation. A restrictor is required in the pressure sensing line for the same reasons as in the case of the relief valve [33, 34]. A complete static and dynamic analysis of a reducing valve has been carried out by Ma [34]. The effect of the restrictor on the stability of the system is emphasised and the analysis supported by experimental results.

Pressure reducing valves are also obtainable in a two stage form, which as in the case of the relief valve is more accurate but has a slower response and is less reliable.

4. CONTINUOUS CONTROL OF HYDRAULIC PRESSURE IN THE L.P. RANGE

4.1 Outline

A design study of a system suitable for continuously controlling pressures in the L.P. range is presented. A method of achieving adequate control with a minimum number of components is described. The system utilises standard components only and is intended to meet the requirements of small scale synthetic diamond manufacture in a reliable and inexpensive manner.

4.2 Definition of the Problem

Pressure control systems for the synthetic diamond process are generally required to be capable of following continuous cycles, the maximum absolute pressure - time slopes of which will not exceed 150 p.s.i./sec. The maximum allowed error at any instant in time will be assumed to be 0.1% of the pressure range used. It is thus sufficient to design the system to satisfactorily follow a ramp function of slope 150 p.s.i./sec.

Other considerations are that the control system should be reliable, inexpensive, and require little maintenance which when necessary may be rapidly performed. All of these factors favour a simple control system which consists of a minimum number of components all of which are readily available.

4.3 Approach to the Problem

The system is envisaged as a single servo-loop as illustrated in Fig. 11. The pump P delivers oil continuously at a fairly constant rate to the cylinder and servo-valve through a check-valve. This part of the circuit is protected by a relief valve which limits the maximum supply pressure to 3,000 p.s.i. The cylinder pressure is sensed by a transducer which converts the pressure signal into a D.C. analogue signal. The reference input is provided by an electronic programmer which generates a D.C. voltage signal in proportion to the required pressure at the particular instant in time. The differential amplifier compares the desired with the actual signal and amplifies the difference. The amplified error activates the electro-hydraulic servo-valve which bypasses more or less oil to the reservoir as required, thus regulating the cylinder pressure.

The system is analysed using the conventional transfer function approach. This requires establishment of the transfer functions of the various elements as a prerequisite to the examination of stability and system response. The elements are individually considered in the following sections.

4.4 The Main Cylinder

In the actual synthetic diamond manufacturing process, the main hydraulic cylinder will be pre-pressurised to a certain value by an auxilliary low-pressure circuit. The purpose of this procedure is to compact the capsule. The pressure limit used varies according to the cylinder and the die-set employed, but is generally of the order 500 p.s.i. in the L.P. range.

In order to obtain the transfer function of the cylinder the following assumptions are made:

(i) The cylinder is at a constant pressure, P_o.

(ii) For pressures in excess of 500 p.s.i. further ram travel is negligible.

(iii) The cylinder volume, V_o is constant and compressibility effects are taken into account in the effective bulk modulus.

(iv) The effective Bulk Modulus, β_e, due to fluid and cylinder is invariant with pressure.

If an incremental volume of fluid, v, is injected into the cylinder, the increment in pressure, p, may be represented by:-

$$p = \frac{\beta_e v}{V_o} \tag{1}$$

Thus the rate of pressure rise, $\frac{dp}{dt}$ may be obtained by differentiating equation (1):

$$\frac{dp}{dt} = \frac{\beta_e dv}{V_o dt} \tag{2}$$

The term $\frac{dv}{dt}$ represents the flow to the cylinder and is denoted by "q". Assuming zero initial conditions, Laplace transformation of equation (2) yields the following expression:-

$$\frac{p}{q} = \frac{1}{S}\frac{\beta_e}{V_o} \tag{3}$$

Where S = Laplace Operator.

Thus if the input to the cylinder is taken to be incremental flow, and the output from it incremental pressure, the cylinder acts as a pure integrator, in series with a perfect amplifier of gain $\frac{\beta_e}{V_o}$.

The cylinder volume V_o, is easily computed from the cylinder geometry.

The effective bulk modulus is dependent on several factors viz:-

(i) The actual bulk modulus of the fluid, β_f.

(ii) The mechanical compliance of the cylinder and other parts of the circuit.

(iii) The effect of trapped air in the system.

Bulk modulus of container (cylinder)

If it is assumed that the mechanical rigidity of the other components in the circuit exceeds that of the cylinder, and the volume of fluid in the circuit is small compared to that within the cylinder, then only the cylinder and its contents need be considered in an estimation of the effective bulk modulus.

Since the piston has been assumed to be axially constrained by the load, the flexibility of the cylinder will be governed by the elasticity of its walls. The radial displacement, u, of the inside of a thick-walled cylinder with outside diameter D_o, and inside diameter D, when subjected to an internal hydrostatic pressure increase, Δp, is given [22] as:

$$u = \frac{D\Delta p}{2E}\left(\frac{{D_o}^2 + D^2}{{D_o}^2 - D^2} + \nu\right) \tag{4}$$

Where E = Young's modulus

ν = Poissons ratio.

If ΔV_c is defined as the change in the cylinder volume, the volumetric compressive strain $\frac{\Delta V_c}{V_o} \simeq \frac{\pi D u}{\pi D^2/4} = \frac{4u}{D}$.

Thus the bulk modulus of the container, β_c is given by:

$$\beta_c = \frac{V_o}{\Delta V_c} \cdot \Delta p = \frac{E}{2} \left(\frac{D_o^2 - D^2}{(1 + \nu)D_o^2 + (1 - \nu)D^2} \right) \tag{5}$$

Bulk modulus of trapped air

Assume the volume of trapped air at the pressure P_o to be V_a. The bulk modulus of the air, β_a, may be defined as:-

$$\beta_a = - V_a \frac{dP_o}{dV_a} \simeq - V_a \frac{\Delta P_o}{\Delta V_a} \tag{6}$$

where ΔP_o = change in pressure,

ΔV_a = change in volume of air.

The bulk modulus will depend on whether isothermal or adiabatic compression of air will take place when pressure is increased. An assumption of adiabatic compression is more realistic [33].

Under adiabatic conditions pressure and volume are related by:-

$$P_o V_a^{\gamma} = \text{constant} \tag{7}$$

where γ = ratio of the specific heats at constant pressure and volume. Differentiating equation (7) the following expression is obtained:-

$$\gamma P_o V_a^{\gamma-1} \, dV_a + V_a^{\gamma} \, dP_o = 0$$

whence
$$\frac{dP_o}{dV_a} = - \frac{\gamma P_o}{V_a} \tag{7.1}$$

Substituting equation (7.1) into equation (6), the adiabatic bulk modulus of air becomes:-

$$\beta_a = \gamma P_o \tag{8}$$

The bulk modulus of the trapped air is directly proportional to the pressure level of the system. Thus one of the advantages of using high pressure systems is that this bulk modulus acquires a greater value and leads to a stiffer system.

The effective bulk modulus

If a change in pressure, ΔP_o in a closed vessel of volume V_o is associated with a change in volume ΔV_o, then:-

$$\Delta V_o = \Delta V_c - \Delta V_f - \Delta V_a \tag{9}$$

where ΔV_c = change in container volume

ΔV_f = change in oil volume

ΔV_a = change in air volume.

The bulk moduli of container, oil and air are respectively defined by:-

$$\beta_c = \frac{V_o \Delta P_o}{\Delta V_c}, \quad \beta_f = -\frac{V_f \Delta P_o}{\Delta V_f} \text{ and } \beta_a = -\frac{V_a \Delta P_o}{\Delta V_a},$$

where V_f = volume of oil in the container.

If the effective bulk modulus is defined as $\beta_e = \dfrac{V_o \Delta P_o}{\Delta V_o}$, then by substituting for ΔV_o, ΔV_c, ΔV_f and ΔV_a in equation (9), and rearranging the following expression is obtained:-

$$\frac{1}{\beta_e} = \frac{1}{\beta_c} + \frac{1}{\beta_f} + \frac{V_a}{V_o}\left(\frac{1}{\beta_a} - \frac{1}{\beta_f}\right) \quad . \tag{10}$$

For systems in the L.P. and I.P. range $\beta_f >> \beta_a$ hence equation (10) may be expressed as:-

$$\frac{1}{\beta_e} = \frac{1}{\beta_c} + \frac{1}{\beta_f} + \frac{V_a}{V_o}\left(\frac{1}{\beta_a}\right). \qquad (10.1)$$

Since this equation involves reciprocals, the effective bulk modulus will be less than any one of the values β_c, β_f or $\frac{V_o}{V_a}\beta_a$.

It is evident that a small amount of entrapped air can drastically reduce the effective bulk modulus, particularly at low pressures.

The amount of trapped air in a system cannot be predicted, thus the bulk modulus can only be estimated. This effect of the bulk modulus on the cylinder gain does not impose any real difficulty, since this may be compensated for elsewhere. A more serious effect of a lowered bulk modulus is the reduction in resonant frequency of actuators in the circuit [see section 4.6]. This can lead to instability of the control system.

Trapped air may be virtually eliminated by careful design of the components and the hydraulic circuit. Blind holes, pockets and tortuous passages should be avoided. In a carefully designed system, virtually all the air will be 'flushed out' by the moving oil shortly after starting up.

4.5 The Pump

Positive-displacement pumps are particularly suitable for this type of application as their delivery is almost independent of back pressure. Although variable displacement type axial plunger pumps are available, the use of a fixed displacement type will incur a substantial saving in cost. A knowledge of the delivery requirements of the pump is however necessary.

Since the pump is arranged to deliver continuously, the maximum rate of pressure rise in the cylinder will occur when the servo-valve is blocked. Under these conditions the flow to the cylinder will be the total delivery of the pump. The required pump output may be predicted by rearranging equation (2), thus:-

$$q_{min} = \frac{V_o}{\beta e}\left(\frac{dp}{dt}\right)_{max} \qquad (2.1)$$

where q_{min} = minimum allowed delivery of the pump.

The required delivery of the pump is dependent on the effective bulk modulus of the system which, it has been shown, cannot be accurately predicted.

If the effective bulk modulus is determined by neglecting trapped air (i.e. $V_a = 0$) and applying equation (10.1), and the minimum pump delivery predicted using equation (2.1), then experience has shown [35] that the delivery of pump selected should be at least three times q_{min}, for circuits of this nature operating in the L.P. range.

If the delivery of the pump required is over-estimated, this implies a small power wastage in a L.P. system. If the delivery of the pump is for example 6 c.i.s.* at a pressure of 3,000 p.s.i., the hydraulic power output is less than 3 HP. This represents a very small proportion of the total power consumption of even a 500 ton capacity synthetic diamond press.

---

* 1 c.i.s. = 1 cubic inch per second.

4.6 The Electro-Hydraulic Servo-Valve

Electro-hydraulic servo-valves convert low power electrical signals into valve-spool motion which may regulate the flow through the valve. They usually have two hydraulic stages, the first of which is an hydraulic pre-amplifier. The electric signal drives a torque motor which actuates the first hydraulic stage. This amplifies the torque-motor force, and is used to drive the second stage.

Due to the two stages of amplification, relatively high flow rates are controllable with low power electric input, and valve operation can be made virtually free of flow forces, stiction and other disturbances. Most two stage servo-valves have a torque motor of the permanent magnet type, and the second hydraulic stage of the four-way-spool valve type.

Commercially available servo-valves differ in the type of first-stage preamplification used, and the feedback employed. The most frequently used means of first stage amplification are jet pipe, nozzle-flapper and spool-valve. Types of feedback generally employed are more numerous and depend upon the control application required and the manufacturer of the servo-valve.

A typical servo-valve, having a nozzle-flapper first stage amplifier and position feedback is schematically illustrated in Fig. 12. Pressurised fluid enters the valve through a single port which distributes it to supply ports 1 and 2. The pressurised fluid enters chambers 1 and 2 via the fixed orifices 1 and 2. If the flapper is centrally positioned, the pressure drops across nozzles 1 and 2 are equal and there is no net force acting on the spool.

If a current is applied to the torque motor, which causes the flapper to move for example towards nozzle 1, the pressure drop across nozzle 1 will decrease while that across nozzle 2 increases. Thus the pressure in chamber 1 increases while that in chamber 2 falls. The pressure difference across the spool will cause the spool to move to the right, thus supplying port 'A' with oil, and simultaneously venting port 'B'. The spool continues to move until the torque on the flapper due to the feedback spring balances the torque due to the input current. At this point the flapper is approximately centred between the nozzles, but the spool has taken a new position directly proportional to the input current. Servo-valves are usually designed with rectangular ports so that the flow through the load ports is directly proportional to spool displacement.

For this particular application several different types of commercially available servo-valves are suitable.

In selecting a particular servo-valve consideration should be given to system flow and response requirements, proven reliability and standardisation with other equipment utilised and maintained by the user.

For the current application a flow control valve with position feedback is considered most suitable. In most valves of this type the main spool quiescent operating point can be adjusted either directly or indirectly. In valves of this type the null-flow of the valve, under zero current conditions can be made equal to the pump output at any chosen pressure.

The flow-input current varies slightly from one make of servo-valve to the other. However the characteristics are generally of the form illustrated in Fig. 13 [33]. Fig. 13 shows that the incremental flow through a typical servo-valve about the null point is directly proportional to the input current up to about 60% to 70% of the rated current. The flow through the valve increases with valve pressure drop. Each curve represents a constant valve pressure drop.

The characteristics presented in the foregoing represent steady state valve behaviour with no output loading; i.e. the output port discharges fluid directly into the reservoir. This resembles the conditions under which the valve is to operate in practice. Although the valve is of the "four-way" construction it is to be used as a three-way valve in this application. With reference to Fig. 12 this may be achieved by permanently blocking port B external to the valve, connecting port A and the return port to the reservoir, and connecting the pressure entry port (which distributes to supply ports 1 and 2) directly to the pump-cylinder-cross*. The servo-valve should be adjusted so that when the coil current is zero, the null flow is equal to the pump delivery at the minimum pressure to be controlled, i.e. 500 p.s.i. The coils should be connected so that a positive current will cause the valve spool to move to the left, preventing the bypass of pump flow to the reservoir and thus raising the system pressure. A negative current will similarly cause a fall in the system pressure.

The flow rating of the valve to be used is determined by the delivery of the pump and the point at which the rated flow - rated current

---

* A cross is a 4 way pipe connection.

characteristic (Fig. 13) ceases to be linear. If for example linearity exists up to 60% of rated current, it is desirable that the valve may be completely closed by this proportion of the current. If the output of the pump is say 6 c.i.s. then the flow rating of the valve should be $\frac{100}{60} \times 6 = 10$ c.i.s. This implies that the valve will only be operated between ±60% of its rated coil current.

Static servo-valve behaviour

Theoretically the valve flow - valve pressure drop characteristics should be represented by an orifice law of the form $Q \propto P^{1/2}$, where Q = flow through the valve and P = valve pressure drop.

In practice however, the characteristics are often found to deviate substantially from the orifice law. This is particularly true for valves having nozzle-flapper preamplification systems. With reference to Fig. 12, if it is assumed that the flapper is moved over to the left, thus increasing the pressure in chamber 1, this increased pressure tends to resist motion of the flapper and thus constitutes a degree of pressure feedback. This feedback is quite separate from the main feedback supplied by the leaf spring. At higher supply pressures the resisting force will also be higher and hence the feedback force is increased. The effect of this feedback is to reduce main spool motion to the right. This effect will also be more significant when lower currents are used. The net result of this effect is that main spool displacement is reduced as the supply pressure is increased, at a given coil current. The decreased valve opening offsets the increased supply pressure somewhat thus making the characteristic curves less steep than the orifice law would suggest.

In many cases [36] the flow through the valve appears to be completely independent of valve pressure drop, after a certain valve pressure drop has been reached. The theoretical characteristics are compared with actual characteristics of a Dowty Moog Series 21 valve, in Fig. 14. It should be emphasised that the pressure-flow characteristics can differ substantially from valve to valve.

For a general case it may be assumed that the flow through the servo-valve, Q, may be expressed as a function of spool position, Y, and valve pressure drop, P, as follows:-

$$Q = g\ (Y,\ P) \tag{11}$$

Equation (11) may be linearised by expanding in a Taylors series about a quiescent point defined by Y_o, P_o, and Q_o, and neglecting second and higher order terms.

Thus $$Q = Q_o + (Y - Y_o)\,\frac{\partial}{\partial Y}\,(Q)_{Y = Y_o} + (P - P_o)\,\frac{\partial}{\partial P}\,(Q)_{P = P_o}$$

Whence $$q_v = k_y y + k_p p \tag{12}$$

Where $q_v = Q - Q_o$ = increment in flow through the valve to the reservoir

$y = Y - Y_o$ = change in spool displacement at quiescent point

$p = P - P_o$ = increment in valve pressure drop

$k_y = \frac{\partial}{\partial Y}\,(Q)_{Y = Y_o}$

$k_p = \frac{\partial}{\partial P}\,(Q)_{P = P_o}$

k_y may be defined as the valve flow gain and k_p as the flow-pressure gain, both at the quiescent point.

q_v represents an increment in flow from valve to reservoir, i.e. an increment in flow away from the cylinder. If q is defined as the increment in flow from the valve towards the cylinder, then by continuity $q = -q_v$.

Thus $\quad q = -k_y y - k_p p \qquad (13)$

The fact that the flow of oil reaches the valve from the cylinder is immaterial since changes in the total flow are being considered here. Under steady state conditions spool displacement is proportional to input current, but the valve is connected so as to reduce the flow to the reservoir when a positive current is applied. Hence a sign inversion exists and $y \propto -i$, where i = a positive change in current. The valve is set so that there is no current through the coils at the quiescent point. i.e. $i = 0$ when $y = y_o$. Thus i represents the total current in the coil.

It follows that $q = k_i i - k_p p \qquad (14)$

where k_i = flow-current gain at the quiescent point.

The input to the valve coils is a potential difference and under steady state conditions is related to the current by Ohm's Law. Thus equation (14) may be written as:-

$$q = k_e e - k_p p \qquad (15)$$

where $\quad k_e = k_i/R$ and R = coil resistance

e = error voltage input to the servo-valve

and $\quad k_e$ is the steady-state flow-voltage gain of the servo-valve.

The value of k_i is the slope of the flow - current characteristic curves, Fig. 13. The coil resistance is supplied by the manufacturer and hence k_e may be determined. k_p is the slope of flow - pressure characteristics, Fig. 14. If the valve is to be operated at a particular point only, k_p may be determined at that point. If the valve is to be operated over a range, as in the current case, k_p should be averaged over the range. If it is assumed that Dowty Moog series 21 type valve is to be used then a value of $k_p = 0$ would be most realistic. In this case equation (15) reduces to

$$q = k_e e \ . \tag{16}$$

Dynamic behaviour of the servo-valve

A complete dynamic analysis of the servo-valve is beyond the scope of this study. Merrit [33] provides a detailed analysis of several different types of servo-valve. He shows that although the transfer functions are usually fourth or sixth order, they are generally dominated by a single lag or quadratic term.

Servo-valve manufacturers usually supply a transfer function with their product. In most cases, the transfer function given is first or second order, and frequency response charts are provided so that the closeness of the fit of the approximation to the actual response may be scrutinised.

For the current application it will be assumed that the servo-valve is a 'Dowty Moog type' whose transfer function may be approximated by the following expression

$$\frac{q}{e} = \frac{k_e}{1 + S\tau_1} \tag{17}$$

where the time constant τ_1 is given by the manufacturer as

$$\tau_1 = 0.0016 \text{ sec} = 1.6 \text{ ms.}$$

Equation (17) reduces to equation (16) under steady state conditions.

4.7 The Pressure Transducer

Since the pressure transducer provides the feedback from output to input, its accuracy governs the maximum accuracy the control system can achieve. It is thus essential that the pressure transducer at least meets the performance specifications demanded for the system. Thus the maximum allowed error on either linearity or reproducibility is 0.1% of full scale pressure transducer output. It is also essential that the transducer be reliable and have a rapid response.

For pressures of up to 3,000 p.s.i., many suitable transducers are commercially available. Ultimate selection will therefore be made on the grounds of proven reliability and cost.

In the author's experience a 0-3,000 p.s.i. bourdon tube-differential transformer type of transducer manufactured by Hartman and Braun has been found to be highly satisfactory.

In this device the pressure is sensed by a bourdon tube with a very small 'dead-volume'. The bourdon tube displaces the iron core of a differential transformer, which is powered by a constant voltage source. The output from the transformer is demodulated and filtered. This output is proportional to the pressure being measured. The transfer function of the pressure transducer may be approximated by a simple lag term thus:-

$$\frac{U}{P_c} = \frac{h}{1 + S\tau_2} \qquad (18)$$

where U = voltage output of the transducer

P_c = cylinder pressure

h = steady-state gain of the transducer

τ_2 = time constant of the transducer

$\simeq \frac{1}{1,000}$ sec. $\simeq$ 1 ms.

The gain, h of the transducer is readily adjustable, while the null point or 'zero' of the device may be shifted by adjusting the initial transfer core position. Thus the transducer may be set to give zero output at 500 p.s.i. cylinder pressure. In this case P_c in equation (18) becomes the increment in cylinder pressure above 500 p.s.i.

4.8 The Programmer

The programmer provides the reference input signal to the control system. In practice the programmer can be of almost any form. Systems used vary from simple mechanically and photo-electrically followed cams to sophisticated electronic programmers, and digital to analogue converters which respond to punched paper tape or magnetic tape. Irrespective of the type of programmer used, it is essential that alterations and adjustments to the programme may be easily and swiftly performed. It is imperative that the programmer be checked independently of the control system for accuracy and reproducibility, and that it be installed only when its reliability has been established.

The differential amplifier which compares the transducer output with that of the programmer and drives the servo-valve may be a standard solid state device. It should have adjustable gain and usually has negligible time constant.

4.9 Analysis of the Control System

The block diagram describing the control system may be constructed using equations (3), (17) and (18). If h is the gain of the pressure transducer and p is the required pressure, then the input, E, to the control system should be E = hp to synchronise programmer and transducer.

The block diagram appears in Fig. 15;

where k_a = gain of the differential amplifier

$k_c = \dfrac{\beta_e}{V_o}$ = cylinder gain

q_d = flow disturbance due to pump, leakages etc.

p_c = cylinder pressure (output)

The open-loop transfer function, GH(S), of the control system is:-

$$GH(S) = \frac{k_a k_e k_c h}{S(1 + S\tau_1)(1 + S\tau_2)} \tag{19}$$

If p is regarded as the input, then the factor h becomes part of the forward path of the block diagram of Fig. 15.

The error response, $\varepsilon(S)$, is defined by:-

$$\frac{\varepsilon(S)}{p(S)} = \frac{1}{1 + GH(S)}$$

thus
$$\frac{\varepsilon(S)}{p(S)} = \frac{S(1 + S\tau_1)(1 + S\tau_2)}{k + S(1 + S\tau_1)(1 + S\tau_2)} \tag{20}$$

where $k = k_a k_e k_c h$.

The steady state error response to the input as a ramp function, i.e. $p(S) = \dfrac{p}{S^2}$, is determined as follows:-

$$\varepsilon = \lim_{S\to 0} S \left[\frac{p}{S^2} \cdot \frac{S(1 + S\tau_1)(1 + S\tau_2)}{k + S(1 + S\tau_1)(1 + S\tau_2)} \right]$$

hence

$$\varepsilon = \frac{p}{k} \tag{21}$$

where ε = steady state error.

This error is required to be less than 0.1% to meet the specifications of the system.

For a maximum pressure of 3,000 p.s.i., ε = 3 p.s.i. Since the maximum pressure ramp to be used is 150 p.s.i./sec., the minimum value of k may be calculated from equation (21):-

Thus $k \geqslant \frac{150}{3}$

$\therefore$ $k \geqslant 50$.

Stability of the system

The characteristic equation of the system is

$$1 + GH(S) = 0.$$

Hence from equation (19),

$$S(1 + S\tau_1)(1 + S\tau_2) + k = 0$$

hence

$$(\tau_1\tau_2)S^3 + (\tau_1 + \tau_2)S^2 + S + k = 0 \tag{22}$$

By applying the Routh-Hurwitz Criterion to equation (22),

The system is stable for $0 \leqslant k \leqslant \frac{\tau_1 + \tau_2}{\tau_1\tau_2}$ (23)

For τ_1 = 1.6 ms., τ_2 = 1 ms., From inequality (23), $0 \leqslant k \leqslant 1630$.

Combining this with the constraint on k imposed by the accuracy requirement,

$$50 \leqslant k \leqslant 1630. \tag{24}$$

A Bode frequency response plot of the open loop transfer function (equation (19)) appears in Fig. 16, for k = 50. The break frequencies, $\frac{1}{\tau_1}$ and $\frac{1}{\tau_2}$ occur at approximately 628 and 1000 radians/sec. respectively. The gain margin is about 30 db. as expected.

The effect of flow disturbances

Changes in the pump output or in leakages in the flow path between the servo-valve and the cylinder constitute flow disturbances. These may be temporary disturbances (pulses) or permanent disturbances (steps).

The response of the cylinder pressure to a disturbance $q_d(S)$ is given by:-

$$\frac{p_c(S)}{q_d(S)} = \frac{k_c(1 + S\tau_1)(1 + S\tau_2)}{S(1 + S\tau_1)(1 + S\tau_2) + k} \tag{25}$$

If the disturbance is an impulse, the steady state response $p_c(t)$ is zero.

If the disturbance is a step, i.e. $q_d(S) = \frac{q_d}{S}$, from equation (25),

$p_c(t)_\infty = \frac{q_d k_c}{k}$.

This is due to the fact that a change in the null operating point is necessary to deal with the changed flow through the servo-valve, and hence a non-zero coil current must exist. A non-zero coil current can only exist if there is a difference between required and actual pressure.

For $k >> k_c$, and q_d small, the steady state error due to flow disturbances will be small.

The transient response of the system

Constraints on the value of the loop-gain, k, have been imposed using required accuracy, and absolute stability (inequality (24)). It is possible to find a value of k, subject to these constraints which optimises the transient response of the system according to some chosen criterion. A method for doing this is described and illustrated in Chapter 6.

If it is assumed that a value of k has been decided upon, the transient response to a ramp input may be determined as follows:-

The closed-loop transfer function is given by:-

$$\frac{p_c(S)}{p(S)} = \frac{k(1 + S\tau_2)}{S(1 + S\tau_1)(1 + S\tau_2) + k} \qquad (26)$$

For p(S) a ramp function $p(S) = \frac{p}{S^2}$

$$p_c(S) = \frac{pk(1 + S\tau_2)}{S^3(1 + S\tau_1)(1 + S\tau_2) + kS^2} \qquad (27)$$

Equation (27) may be written in the form:-

$$p_c(S) = \frac{pk}{\tau_1} \cdot \frac{\frac{1}{\tau_2} + S}{S^2(S + \alpha_1)(S + \alpha_2)(S + \alpha_3)} \qquad (28)$$

where α_1, α_2, α_3 are constants.

The inverse transform is of the form:-

$$p_c(t) = pt - \frac{p}{k} + p(C_1 e^{-\alpha_1 t} + C_2 e^{-\alpha_2 t} + C_3 e^{-\alpha_3 t}) \qquad (29)$$

where C_1, C_2, C_3 are constants, and $C_1 + C_2 + C_3 = \frac{1}{k}$.

After the transient terms have decayed the response will be given by

$$p_c = pt - \frac{p}{k} .$$

The expected response is illustrated in Fig. 17.

5. CONTINUOUS CONTROL OF HYDRAULIC PRESSURE IN THE I.P. RANGE

5.1 Outline

A means of extending the working pressure of the basic control system discussed in Chapter 4, into the I.P. range is described. The method involves the incorporation of a non-standard pressure reducing device into the circuit. The device, which has been named a "Proportional Pressure Divider" is analysed in a design study.

5.2 Definition of the Problem

For large scale synthetic diamond manufacture, cylinder pressures in the I.P. range are required (see section 2.7). For the present case it will be assumed that the maximum cylinder pressure to be used is 10,000 p.s.i., and that a pressure control system, capable of following a continuous pressure - time cycle in the range 1,000 p.s.i. to 10,000 p.s.i. is required. The limit of 150 p.s.i./sec. on pressure rise and fall rates, and the allowed error of 0.1% of the maximum pressure used in Chapter 4, will be assumed in this case too.

The use of a 'scaled-up' version of the control system discussed in Chapter 4 is prevented by the unavailability of servo-valves capable of operating at the pressure levels required.

This is due primarily to the fact that electro-hydraulic servo-valves were originally developed for use in aircraft and for aero-space purposes. The working pressure for the hydraulic systems employed for these applications is determined on the grounds of minimum weight of the control system, pumps, actuators etc. The optimum working pressure for almost all systems of this type falls between 3,000 and 4,000 p.s.i.[33, 37].

For industrial purposes, where working pressure is chosen on the grounds of minimum cost, pressures in excess of 3,000 p.s.i. are not often used. Thus commercially available servo-valves are at present limited to working pressures of under 4,000 p.s.i.

5.3 Approach to the Problem

One possible approach is to design and develop a servo-valve capable of operating at the required pressure. The expense of a research and development programme, and the cost of tooling etc. to produce a satisfactory device is only justified if a large market exists for higher pressure servo-valves. At present this is not the case.

Another approach is to incorporate a pressure reducing valve between the cylinder and the servo-valve. When utilised in this manner, the reducing valve merely protects the servo-valve from the I.P.

It is essential that the reducing valve does not interfere with the flow requirements of the servo-valve. In the circuit, pressure is controlled by regulating the fluid by-passed through the servo-valve. If at any time the servo-valve is 'starved', loss of control over the cylinder pressure will result. It is also essential that the reducing valve be reliable and stable over the whole pressure range.

Standard pressure reducing valves are not designed to meet the requirements set out above. They are generally intended to work at a constant supply pressure and can become unstable if the supply pressure varies significantly. They do not 'fail-safe' if the stabilising orifice becomes blocked, and thus a relief valve would have to be used in the circuit to safeguard the servo-valve against this eventuality. Another difficulty is the fact that standard reducing valves do not generally respond sufficiently rapidly to prevent 'starving' of the servo-valve.

A pressure reducing valve capable of meeting the requiremnts discussed in the foregoing is described in reference [38]. The device has been named a "Proportional Pressure Divider" since it reduces the supply pressure by a constant ratio.

The incorporation of the reducing valve into the basic control system is schematically illustrated in Fig. 18. The proportional pressure divider is described and analysed in the following sections.

5.4 Description of the Proportional Pressure Divider (PPD)

For this particular case the reduction factor required is 0.3 . (From a maximum pressure of 10,000 p.s.i. cylinder or primary pressure to a maximum of 3,000 p.s.i. servo-valve or secondary pressure). The proportional pressure divider is illustrated in Fig. 19.

It consists of two pistons operating in two bores in the same cylinder block. The ratio of the areas of the two pistons is the same as the pressure reduction required, (in this case 10:3). The smaller, or primary piston operates in the higher pressure region of the valve, and has a needle valve at its one end, which is lapped into a seat. The secondary piston operates in the reduced pressure zone and has a damper at its one end. The two pistons butt up against each other.

I.P. fluid from the cylinder of the press P_1 is introduced at the inlet port. If $P_2 < \frac{3}{10} P_1$ the force differential across the pistons causes the needle valve to open and oil flows through to the secondary side until such time as P_2 just exceeds $\frac{3}{10} P_1$, at which stage the needle valve closes. If on the other hand, $P_2 > \frac{3}{10} P_1$, the needle valve will remain shut until the servo-valve, which is connected to P_2, bleeds away the excess pressure in the secondary side. Control of primary pressure is thus achieved by regulating the pressure in the secondary using the servo-valve.

Pressure reduction actually takes place in the throttle formed by the needle valve and seat. Feedback action is by force comparison and the system is highly stable.

To ensure that no build up of back-pressure takes place behind the pistons, the area is vented by a connection to the reservoir. Any leakage past the pistons reaches the reservoir through this path.

5.5 Analysis of the PPD

The cylinder pressure adjustment is effected by regulating the flow using the servo-valve. It is thus important that the PPD does not interfere with the flow demand of the servo-valve.

The PPD can affect the flow in three possible ways, namely:-

(i) There is a natural slight increase in flow through the PPD since the oil is less compressed due to lower pressure in the secondary side.

(ii) There is a flow loss due to leakage past the pistons.

(iii) If the spool response is too slow, the needle valve of the PPD rather than the servo-valve can limit the flow.

The first two effects partially cancel each other.

Leakage past the pistons is easily minimised by ensuring that the pistons are close fits in their respective cylinders, and are made sufficiently long. Annular oil grooves on the pistons also reduce leakage losses and ensure adequate lubrication.

The third effect can be minimised by ensuring that the PPD responds sufficiently rapidly to meet the flow requirements of the servo-valve.

Fig. 20 is a block diagram representation of the PPD. The pressure drop across the throttle is dependant upon the flow through the throttle and the needle valve opening. A unique valve opening will exist for each different flow condition at which the pressure drop across the throttle is consistent with the required pressure ratio.

When the needle valve is in this position the throttle behaves as a pressure attenuator with constant gain. Any change in flow, primary or secondary pressure will upset the equilibrium. This gives rise to a differential force which acts upon the pistons repositioning them so as to restore equilibrium.

The flow through the throttle is related to the pressure difference and the valve opening by the following non-linear expression:-

$$Q = CA(\chi)(P_1 - P_r)^{1/2} \qquad (30)$$

where Q = flow through the throttle

C = a dimensional constant

$A(\chi)$ = effective orifice area, a function of valve opening.

P_1 = primary pressure

P_r = reduced pressure behind the throttle

χ = valve opening

Equation (30) may be linearised about a quiescent operating point defined by Q_o, χ_o, P_{1o}, P_{ro}.

If q is defined as the increment in flow

p_1 is defined as the increment in primary pressure

p_r is defined as the increment in reduced pressure

and x is defined as the increment in valve opening,

then

$$q = \left(\frac{\partial}{\partial \chi} A(\chi)\right)_{\chi=\chi_o} \cdot \frac{Q_o}{A(\chi_o)} \cdot x + \frac{Q_o(p_1 - p_r)}{2(P_{1o} - P_{ro})}$$

hence

$$q = k_1 x + k_2(p_1 - p_r)$$

and finally

$$p_r = \frac{k_1}{k_2} x + p_1 - \frac{q}{k_2} \tag{31}$$

where $k_1 = \dfrac{Q_o}{A(\chi_o)} \cdot \left(\dfrac{\partial A(\chi)}{\partial \chi}\right)_{\chi=\chi_o}$ and $k_2 = \dfrac{Q_o}{2P_{1o} - 2P_{ro}}$

k_1 may be defined as the valve flow-displacement gain and k_2 as the valve flow-pressure gain. k_1 and k_2 are approximately constant only for small disturbances about the quiescent point.

The frictional resistance of the flow path from behind the throttle to the secondary chamber, and the equivalent capacitance due to the compressibility of the oil in this 'closed' volume cause a lag while the secondary pressure builds up.

If the resistance of the path is denoted R_1, and the flow through it q_2, then for small pressure differences the flow is laminar.

Thus $$p_r - p_2 = R_1 q_2 \tag{32}$$

where p_2 = incremental secondary pressure consistent with the foregoing analysis.

But $$q_2 = \frac{V_2}{\beta_e} \frac{dp_2}{dt} \tag{33}$$

where V_2 = total enclosed volume of flow path and secondary piston chamber

and β_e = effective bulk modulus of fluid.

Thus

$$p_r - p_2 = \frac{R_1 V_2}{\beta_e} . \frac{dp_2}{dt} \tag{34}$$

If ω_2 is defined as

$\frac{\beta_e}{R_1 V_2}$ = hydraulic natural frequency of the secondary flow path and chamber,

then by Laplace transforming equation (34) and rearranging the following expression is obtained:-

$$\frac{p_2}{p_r} = \frac{1}{1 + \frac{S}{\omega_2}} \tag{35}$$

Due to the pressure drop across the needle valve throttle, a small force exists which tends to close the valve. If this force is ignored as a small disturbance, then the following expression describes the piston dynamics:-

$$p_1A_1 - p_2A_2 = (MS^2 + BS)\,x$$

i.e.

$$x = \frac{p_1A_1 - p_2A_2}{MS^2 + BS} \qquad (36)$$

where A_1 = primary piston area

A_2 = secondary piston area

M = mass of pistons and fluid in the chambers

B = damping constant.

Equations (31), (35) and (36) may be used to establish the relationships in the block diagram, Fig. 20.

The various transfer functions of interest are readily obtainable from the block diagram.

5.6 Steady-State Pressure Response

From the block diagram, Fig. 20:-

$$\frac{p_2}{p_1} = \frac{S(S + \frac{B}{M}) + \frac{A_1k_1}{Mk_2}}{S(S + \frac{B}{M})(1 + \frac{S}{\omega_2}) + \frac{A_2k_1}{Mk_2}} \qquad (37)$$

Thus the steady-state response to step input increment, $\frac{p_1}{S}$ is:-

$$p_2 = p_1 \frac{A_1}{A_2} \qquad (38)$$

which is the desired pressure ratio. The fact that k_1 and k_2 vary over the operating range does not affect the steady-state pressure ratio.

5.7 Stability

The open-loop transfer function of the system is given by:-

$$GH(S) = \frac{k_s}{S(1 + \frac{S}{\omega_2})(1 + \frac{S}{\omega_s})} \tag{39}$$

where $k_s = \frac{A_2 k_1}{B k_2}$

and $\omega_s = \frac{B}{M}$ = piston mass - damper natural frequency.

The polar frequency response plot of equation (39) appears in Fig. 21 (a). This system is stable if k_s is not too large. However, since k_s depends on quantities that can vary considerably over the working range of the valve the possibility of instability cannot be ruled out.

The lag term $(1 + \frac{S}{\omega_2})$ due to the flow path resistance can be eliminated by ensuring that $\frac{1}{\omega_2}$ is very small. In practice this is achievable if the flow path is unrestricted and the dead volume of the secondary chamber is kept low.

Under such conditions, equation (39) may be rewritten thus:-

$$GH(S) \simeq \frac{k_s}{S(1 + \frac{S}{\omega_s})} \tag{40}$$

The polar plot of equation (40) approaches the inherently stable form illustrated in Fig. 21 (b), and stability is assured.

5.8 Transient Pressure Response

If the lag term $(1 + \frac{S}{\omega_2})$ is ignored, equation (37) may be rewritten as follows:-

$$\frac{p_2}{p_1} = 1 - \frac{\omega_n^{\ 2}\left(\dfrac{A_2 - A_1}{A_2}\right)}{S^2 + 2\zeta\omega_n S + \omega_n^{\ 2}} \tag{41}$$

where $\omega_n = \sqrt{\dfrac{A_2 k_1}{k_2 M}}$ is the natural frequency

and $\zeta = \dfrac{B}{2M\omega_n} = \dfrac{B}{2M}\sqrt{\dfrac{k_2 M}{A_2 k_1}}$ is the damping ratio. (42)

The response to a step $p_1(S) = \dfrac{P_1}{S}$ is:-

$$p_2(S) = \frac{P_1}{S} - \frac{P_1\omega_n^{\ 2}\left(\dfrac{A_2 - A_1}{A_2}\right)}{S(S^2 + 2\zeta\omega_n S + \omega_n^{\ 2})}$$

hence

$$p_2(t) = \frac{P_1 A_1}{A_2}\left[1 - \frac{e^{-\zeta\omega_n t}}{\sqrt{1-\zeta^2}}\sin(\omega_n t + \theta)\right] \tag{43}$$

where $\theta = \tan^{-1}\dfrac{\sqrt{1-\zeta^2}}{\zeta}$.

The transient pressure response may be optimised according to some criterion using a method detailed in Chapter 6. The optimisation would be carried using ζ as the variable. The dynamic behaviour of the system is limited by the value of the natural frequency, ω_n; it is thus desirable to make ω_n as high as possible.

The natural frequency can be made high by minimising the piston mass and ensuring that the flow-displacement gain k_1 is large as compared with the flow-pressure gain k_2. This may be effected by adjusting the angle of the needle valve.

Equation (43) suggests that the transient pressure response will be that of a second order system. The response is illustrated in Fig. 22 using an assumed damping ratio of $\zeta = 0.7$.
The percentage overshoot [39] is $100e^{-\zeta\pi/(1 - \zeta^2)} \simeq 5\%$ and the setting time [39] is $\dfrac{4}{\zeta\omega_n} \simeq \dfrac{5.7}{\omega_n}$.

Once ω_n has been established, the damping constant B may be determined from equation (42) to give ζ the value determined in the optimisation analysis. There may be considerable variation in both the natural frequency and the damping ratio due to the variations of k_1 and k_2. The natural frequency and damping ratio should be averaged over the operating range of the PPD.

5.9 Damper Design

With reference to Fig. 23, neglecting compressibility effects the flow, Q, past the plunger is given by [40] as:-

$$Q = \frac{\pi}{12} \frac{Dd^3 \Delta p}{L\mu} \tag{44}$$

where D = plunger diameter

d = mean clearance between plunger and bore

A = net piston area

Δp = pressure drop across piston

L = length of plunger

μ = fluid viscosity.

The damping force is

$$A\Delta p = B\frac{dx}{dt} \tag{45}$$

and since $\frac{dx}{dt} = \frac{Q}{A}$ it follows that

$$B = \frac{12A^2L\mu}{\pi Dd^3} \tag{46}$$

By making the plunger a good fit in the bore at first, the damping constant may be altered by changing the clearance or drilling holes through the piston.

5.10 Response of Secondary Pressure to a Flow Disturbance

From the block diagram, Fig. 20, ignoring the lag $\frac{1}{\frac{S}{\omega_2} + 1}$,

$$\frac{p_2}{q} = \frac{-BS\left(\frac{S}{\omega_s} + 1\right)}{k_1A_2 + k_2BS\left(\frac{S}{\omega_s} + 1\right)} \tag{47}$$

thus the steady-state pressure response to a step change in flow is zero. The steady-state pressure response to a displacement disturbance is similarly zero.

5.11 Piston Response to Flow Changes

From the block diagram, Fig. 20, ignoring the lag $\frac{1}{\frac{S}{\omega_2} + 1}$,

$$\frac{x}{q_2} = \frac{\frac{\omega_n^2}{k_1}}{S^2 + 2\zeta\omega_nS + \omega_n^2} \tag{48}$$

where ζ and ω_n are as previously defined.

Thus for a unit step change in the flow, i.e. $q_2(S) = \frac{1}{S}$

$$x(t) = \frac{1}{k_1} \left[1 - \frac{e^{-\zeta\omega_n t}}{\sqrt{1 - \zeta^2}} \sin (\omega_n t + \theta) \right] \tag{49}$$

The steady-state change in the displacement is thus $\frac{1}{k_1}$, the inverse of the flow-displacement gain.

The damping ratio, percentage overshoot and settling time are the same as in the case of the transient pressure response. If the optimum value of the damping ratio is used in conjunction with a high natural frequency, a rapid piston response is achievable.

5.12 The Effect of the PPD on the Overall Control System

It was specified in section 5.2 that the control system should be capable of controlling cylinder pressure in the range 1,000 p.s.i. to 10,000 p.s.i. Most servo-valves suffer a severe loss in performance if their supply pressure is allowed to drop below about 500 p.s.i. [33, 36]. Most manufacturers recommend that their servo-valves be operated at pressures in excess of this value.

With the PPD in the control circuit, when the servo-valve pressure falls to 500 p.s.i., the cylinder pressure will be about 1,670 p.s.i., and thus difficulty may be experienced in regulating cylinder pressures between 1,000 p.s.i. and 1,670 p.s.i. This difficulty may easily be overcome if an unloading valve is used to bypass the needle valve throttle of the PPD, when the cylinder pressure falls into the L.P. range. The minimum 'unloading pressure'* to which the unloading valve should be set in this case is 1,670 p.s.i. It would however be more beneficial to set the 'unloading pressure' to a value near 3,000 p.s.i., (e.g. 2,800 p.s.i.) so that the servo-valve pressure is at no stage less than about 840 p.s.i. As soon as unloading occurs, the servo-valve pressure will be equal to the cylinder pressure, and will consequently remain above 1,000 p.s.i.

The unloading valve should be arranged to operate on a signal from the secondary pressure side of the PPD rather than the primary. This would have the effect of eliminating any time lag between primary and secondary response. Standard unloading valves are usually fail-safe devices, thus the servo-valve will be protected should the unloading valve malfunction. An unloading valve may easily be incorporated in the construction of the PPD. Numerous unloading valve configurations, some of which may lead to an insignificant addition to the cost of the PPD, are possible.

Inclusion of the PPD into a basic control system of the type discussed in Chapter 4 should not cause any loss in dynamic performance.

The servo-valve is 'starved' only if the pressure in the secondary side drops below 500 p.s.i. Even without an unloading valve to prevent this eventuality, the secondary pressure would not be expected to drop as low as this (for primary pressures in excess of about 2,750 p.s.i.). This is due to the fact that the transfer function of the pressure response (equation (41)) is second order. A natural frequency, ω_n, in excess of

---

* unloading pressure or 'cracking' pressure is the pilot pressure at which a given circuit is bypassed, or unloaded.

1,400 Hz. is easily obtainable. For purposes of comparison with the servo-valve response, the 10% to 90% rise time of the secondary pressure to a step change in primary pressure may be estimated from the half power frequency bandwidth of a second order system [41]. The 10% to 90% rise time, T_r, is given [42] as:-

$$T_r = \frac{2.2}{\omega_p} \tag{50}$$

where ω_p = frequency at which input power is reduced to half.

(i.e. Signal magnitude is attenuated by a factor of $\frac{1}{\sqrt{2}}$, or reduced by 3 db.)

If the damping ratio, ξ, is assumed to be 0.7, then the half-power frequency of a second order transfer function is approximately equal to the natural frequency, i.e. $\omega_p = \omega_n$. Hence from equation (50), $T_r = \frac{2.2}{2\pi \times 1400} = 0.25$ ms. The servo-valve on the other hand is approximated by a first order transfer function (equation (17)), with a time constant of 1.6 ms, which implies a 10% to 90% rise time in excess of 3.5 ms.

Thus the servo-valve spool motion is likely to be considerably slower than the rate at which pressure in the secondary chamber may adjust to prevent 'starving' of the servo-valve.

With the inclusion of an unloading valve in the PPD, the dynamic performance of the overall control system can actually be improved in the L.P. range. For the type of servo-valve chosen in Chapter 4, examination of Fig. 14 reveals that load flow is substantially independent of servo-valve supply pressure when the valve pressure drop (all of the supply pressure is dropped across the valve in this case) exceeds 850 p.s.i. If the unloading valve is arranged to respond to secondary pressure, the secondary pressure should never drop below this value. This will improve the performance of the servo-valve. The design procedure used in Chapter 4 may thus be applied with equal confidence for I.P. circuits which incorporate a PPD.

6. CONTINUOUS CONTROL OF HYDRAULIC PRESSURE IN THE H.P. RANGE

6.1 Outline

The analysis and design of a system suitable for the continuous control of hydraulic pressure in the H.P. range is considered. The system is intended for utilisation in isostatic compaction processes which require predetermined pressure - time cycles, but may in principle be adapted to other applications as well.

With the exception of the hydraulic power amplifier which is not commercially available, and the high pressure chamber, all elements in the system are standard electronic, electro-hydraulic or hydraulic components. The various elements to be used are individually considered, and the design of the high pressure chamber and the hydraulic power amplifier is discussed.

Stability criteria are established and subsequently used as constraints in the optimisation of the control system. The system is optimised according to a certain criterion using a technique which is applicable to a wide variety of control systems. The optimisation is carried out on a digital computer, which is also used to determine the sensitivity of the system to variations in some of the parameters. The transient response of the system to various input functions is simulated using a digital computer. The relative stability of the system is examined using a digital computer simulation of the frequency response.

The computer methods used are generally applicable to other control systems, although certain changes in the programmes used will have to be made in some cases.

6.2 Definition of the Problem

The isostatic compaction process has been described in section 2.8, where it was pointed out that anvils and dies for synthetic diamond manufacture may benefit from being subjected at this stage of the manufacturing procedure, to a predetermined pressure - time cycle. The actual process to be used in this case is further discussed in section 6.5.

Complete isostatic presses capable of operating at pressures of up to 100,000 p.s.i. are commercially available. None of these however provide the facilities for following predetermined continuous pressure - time cycles.

The maximum pressure to be used in this case is 50,000 p.s.i., and a system for generating and controlling pressures up to this level is required. The system should be capable of following a pressure - time programme which may require absolute rise and fall rates as high as 3,000 p.s.i./sec., with minimum deviation. Unlike the synthetic diamond process where the hydraulic power used for pressure control constitutes an insignificant proportion of the total power requirements of the process, in isostatic pressing most of the power used is that required to pressurise the vessel. For pressures of the high order to be used here, even relatively small flow rates result in substantial hydraulic power consumption. The efficient use of hydraulic power is therefore essential. Although simplicity, which fosters low initial cost and reliability is still important, it may be traded off in this case where significant gains in performance or operating efficiency may be achieved. The system should be amenable to mass production procedures; furthermore it is preferable if the system can be easily modified to accommodate a change in operating power which might be required in different applications.

It is considered desirable that standard components be utilised wherever possible, and that the system be sufficiently independent of the pressure vessel so as to facilitate rapid re-attachment to another press.

6.3 Approach to the Problem

The design of a high pressure hydraulic system for an isostatic press requires the solution of two inter-connected problems, namely pressure generation and pressure control. Since pumps capable of producing pressures in the H.P. range are not readily available, high pressure must generally be obtained by the intensification of a lower pressure source. Hence the pressure generation problem reduces to one of pressure intensification.

Although an indefinite number of solutions to the control problem may be proposed, only two basically different approaches to the pressure intensification problem exist. These are internal pressure intensification and external pressure intensification. In the first approach, pressure would be generated within the high pressure chamber by means of a plunger which is driven by a lower pressure source, in much the same manner as was described in connection with organic chemical synthesis (section 2.3) and hydrostatic extrusion (section 2.4). In the second approach, the working fluid would be pressurised external to the pressure vessel and metered to the chamber by the control system as required. The two philosophies of approach to the pressure intensification problem give rise to rather different control problems and should be considered separately.

6.3-1 Internal pressure intensification

A schematic representation of one possible form of isostatic press with internal pressure intensification appears in Fig. 24. The operation of the control system is essentially the same as that described in Chapter 4. The servo-valve which is actuated by the amplified difference between the required and actual high pressure, causes a change in pressure in the L.P. (or I.P.) chamber. The lower pressure change is intensified by the plunger and the resulting change in high pressure reduces the original discrepancy between actual and required values. The system is stabilised by means of a viscous damper.

Although isostatic presses with internal pressure intensification represent workable systems, they possess certain disadvantages. Some of these are considered below:-

(i) Loss of fluid because of leakage from the high pressure chamber through the seals cannot be made up during the H.P. cycle.

Attempts to minimise such leakage may result in excessive seal friction which in turn could severely limit the dynamic performance of the system. Additionally it may result in an inordinate 'dead-zone' in the high pressure response.

It might appear to be possible to circumvent the leakage problem by having a very large initial volume of fluid in the H.P. chamber. But this would result in long plunger strokes which would give rise to an unnecessarily high overall power consumption (see section 6.5) and low natural frequency of the L.P. (I.P.) chamber.

(ii) The intensification system is an integral part of the press, thus modification to accommodate a change in desired operating power might require major alterations.

(iii) Irrespective of its shape, the plunger interferes with the working volume of the high pressure chamber. In some cases special precautions will have to be adopted to ensure that the plunger does not foul the container of the material being compacted.

(iv) The overall system will be bulky and expensive to manufacture.

Most of these disadvantages can be overcome, or minimised by using external pressure intensification.

6.3-2 External pressure intensification

A relatively small pressure intensifier which is arranged to operate continuously externally to the press, may be used to charge an accumulator with high pressure fluid. The high pressure chamber of the isostatic press may be fed from the accumulator when an increase in pressure is desired, or may be bled of some of its fluid when the pressure is to be reduced. When used in this manner the intensifier cannot affect the dynamic behaviour of the control system, or interfere in any way with the working volume of the high pressure cavity. Should a change in operating power be deemed necessary, an additional intensifier and/or accumulator may be added with little inconvenience. Pressure intensifiers and accumulators suitable for operation in the H.P. range are commercially available.

On the strength of the foregoing arguments the external pressure intensification approach is adopted. The design of a suitable high pressure system is discussed in the following sections.

6.4 Description of the Proposed System

The proposed system is schematically illustrated in Fig. 25. The programmer generates a voltage signal proportional to the instantaneous required chamber pressure. The chamber pressure is sensed by the pressure transducer T_1, and the difference between the required and actual pressure is amplified by the differential amplifier. The error signal is modified by the adaptive gain compensator, and used to actuate the electro-hydraulic servo-valve. The servo-valve drives the hydraulic power amplifier which supplies or bleeds fluid from the high pressure chamber as required.

A single I.P. pump supplies the circuit with its fluid power requirements. Fluid reaches the servo-valve via a pressure reducing valve, which since the supply pressure is constant may be of a standard type. The bulk of the pump output is taken by the intensifier, which boosts the pressure and charges the high pressure accumulator. The accumulator acts as a 'buffer' in that it is capable of supplying additional power during peak load periods while it stores power during slack portions of the cycle. When the accumulator is fully charged the supply to the intensifier is cut off by a solenoid operated 'on-off' valve which is actuated by a volume-sensing limit switch in the accumulator. The accumulator feeds the hydraulic power amplifier.

The hydraulic power amplifier (HPA) is intended to keep the high pressure output flow proportional to the low pressure input flow from the servo-valve. The HPA is envisaged as a spool-type valve the output flow of which is dependent on the spool displacement and the pressure drop across it. Proportionality between input flow and valve opening is assured by feeding the spool displacement back to the servo-valve using a position transducer. As the valve pressure drop varies in sympathy with the changing pressure in the high pressure chamber, it alters the gain of HPA. This affects the loop gain of the system which should be kept steady for optimum performance. (See section 6.14).

Variations in the loop gain are compensated for by the adaptive gain compensator which makes use of chamber and accumulator pressure information supplied by transducers T_1 and T_2 respectively, and spool positional information supplied by the displacement transducer, to effect an appropriate adjustment. (See section 6.10).

6.5 Preliminary Considerations

Before the various elements of importance in the control system are examined in detail, certain aspects of the process being controlled, which are of prime importance to the design of the high pressure system, should be considered.

In the actual process, when the flexible container has been charged with material to be compacted, it is subjected to a low amplitude, 50 Hz vibration while the air is being evacuated. This assists air removal. After sealing, the charged container is subjected to a pre-compaction pressure - time cycle which uses a maximum pressure of about 10,000 p.s.i., at which stage the compact has attained about 97% of its final 'green density'*. Thus prepared, the material is ready for the vital H.P. cycle.

It is desirable that the pre-compaction be performed separately from the H.P. compaction cycle to minimise the power requirements of the process. During the pre-compaction cycle, the volume of the material being compacted may be reduced by as much as 65%. Thus after pre-compaction, at least 65% of the high pressure chamber volume would be occupied by hydraulic fluid. Thus most of the work done in raising the pressure within the chamber would be absorbed by the 'dead volume' of fluid. In order to realise the high rates of pressure rise which are desired, the expenditure of enormous quantities of power would be required.

The foregoing may be illustrated by considering a typical example:-

Assume the powder is enclosed in a cylindrical container having internal dimensions of 11.5 in. diameter by 11.5 in. high. The initial volume of powder is thus approximately 1,110 cu. in. After pre-compaction in which it was subjected to pressures of up to 10,000 p.s.i., the material has been reduced to a pellet 8 in. in diameter and 8 in. high, a reduction of volume of about 63.7% to 402 cu. in. The density of the pellet increases by a further 3% of its final density during the H.P.

---

* 'Green density' implies density before sintering.

cycle. The volume of the pellet is thus decreased by a further 12.7 cu. in. to about 389.3 cu. in.

If the volume of the chamber is 1,200 cu. in., the maximum fluid volume in the chamber will be approximately 810 cu. in.

From equation (2) (section 4.4), the rate of pressure rise is given by:-

$$\frac{dp}{dt} = k_c q \tag{51}$$

where $\frac{dp}{dt}$ = rate of change of chamber pressure.

$$k_c = \frac{\beta_e}{V_o} = \text{chamber gain} \tag{52}$$

β_e = effective bulk modulus

V_o = volume of the chamber

q = volume - flow rate of fluid into the chamber.

The hydraulic horse power, W, required to maintain a flow, q, against a pressure P_c is given by:-

$$W = \frac{P_c q}{6600} \tag{53}$$

where W is in horse power, q in c.i.s. and P_c in p.s.i.

From equations (51) and (53),

$$W = \frac{P_c}{6600 k_c} \cdot \frac{dp}{dt} \tag{54}$$

The chamber gain, k_c, may be estimated by using a development similar to that used in section 4.4 . The effect of the compressibility of the material being compacted must however be taken into account.

If a change in chamber pressure, ΔP_c is associated with a total change of ΔV_o in volume, then in the absence of trapped air in the system;

$$\Delta V_o = \Delta V_c - \Delta V_f - \Delta V_m \tag{55}$$

where ΔV_c = change in chamber volume

ΔV_f = change in fluid volume

ΔV_m = change in volume of material being compacted.

The effective bulk modulus, and the bulk moduli of the chamber, fluid, and material are respectively defined as:-

$$\beta_e = \frac{V_o \Delta P_c}{\Delta V_o}\ , \quad \beta_c = \frac{V_o \Delta P_c}{\Delta V_c}\ , \quad \beta_f = -\frac{V_f \Delta P_c}{\Delta V_f}\ , \text{ and } \beta_m = -\frac{V_m \Delta P_c}{\Delta V_m}$$

where V_f = initial volume of fluid

V_m = initial volume of material being compacted.

By substituting for ΔV_o, ΔV_c, ΔV_f, and ΔV_m in equation (55), the following expression is obtained:-

$$\frac{V_o}{\beta_e} = \frac{V_o}{\beta_c} + \frac{V_f}{\beta_f} + \frac{V_m}{\beta_m}$$

Re-arrangement of the above expression yields the following:-

$$\frac{\beta_e}{V_o} = k_c = \frac{\beta_c \beta_f \beta_m}{V_o \beta_f \beta_m + V_f \beta_m \beta_c + V_m \beta_f \beta_c} \tag{56}$$

In the actual process, the pellet of material being compacted approaches its final 'green' density at an early stage in the H.P. cycle. The rest of the cycle is intended to alter the mechanical properties of the compact rather than to increase its density. (The pellet does however suffer relatively small permanent deformations during the rest of the cycle, which may be neglected for the present case since they affect the pellet volume very slightly.) For most of the H.P. cycle the pellet behaves as an elastic solid with an apparent bulk modulus, β_m, of the order 3×10^6 p.s.i.

If the bulk modulus of the chamber is assumed to be of the order 15×10^6 p.s.i., and that of the fluid 3×10^5 p.s.i., then for the current example, where $V_o = 1200$, $V_m = 389.3$, and $V_f = 810$ cu. in.;

$$V_o\beta_f\beta_m \simeq 1.08 \times 10^{14}$$

$$V_f\beta_m\beta_c \simeq 3.65 \times 10^{16}$$

$$V_m\beta_f\beta_c \simeq 1.75 \times 10^{15}$$

and $$\beta_m\beta_f\beta_c \simeq 1.35 \times 10^{19}$$

From equation (56), $k_c \simeq 350\ \mathrm{lb_f.\ in.^{-5}}$

From equation (54), to realise a rate of pressure rise of 3,000 p.s.i./sec when the chamber pressure is 45,000 p.s.i. requires an hydraulic horse power input of approximately 58.5 hp.

This high power requirement is due to the low value of the chamber gain, k_c. The latter may be substantially increased by reducing the volume of fluid, V_f.

Thus if the chamber volume, V_o, is reduced to approximately 440 cu. in. then for a final pellet volume, V_m, of 390 cu. in. the fluid volume V_f will be 50 cu. in.

Thus $$V_o\beta_f\beta_m \simeq 4.00 \times 10^{14}$$

$$V_f\beta_m\beta_c \simeq 2.25 \times 10^{15}$$

$$V_m\beta_f\beta_c \simeq 1.75 \times 10^{15}$$

while $$\beta_m\beta_f\beta_c \simeq 1.35 \times 10^{19} \quad \text{as before.}$$

Applying equation (56), $k_c \simeq 3060$.

The hydraulic power requirements would be correspondingly reduced to about 6.7 hp.

This substantial reduction in required power, and the associated reduction in the size of the pump, intensifier, accumulator, and other components, well justifies the division of the compaction process into two separate cycles.

The pressures required during the pre-compaction cycle are all in the I.P. range, hence a control system similar to that described in Chapter 5 may be used. The absolute rates of pressure rise and fall required in this portion of the cycle are relatively low, (of the order 50 p.s.i./sec maximum). This is fortunate since the material being compacted is highly compressible for a large part of the cycle, and this will result in a low chamber gain.

The two compaction cycles may be carried out using the same basic press, although separate presses may render the process more amenable to mass production.

If the same press is to be used, the high pressure chamber should be sufficiently large to accommodate the charged container prior to the pre-compaction cycle. For the purpose of carrying out the H.P. compaction cycle, a highly incompressible 'blank' in the form of a hollow cylinder with one end open could be slid into the chamber to take up the dead volume. The arrangement is illustrated in Fig. 26. The 'cavity' of the 'blank' would be sufficiently large to accommodate the pre-compacted pellet, which at this stage has been removed from its original container and coated with a thin covering of flexible plastic material. The blank should be a loose fit in the high pressure cavity, and several small holes should be drilled through its bottom and walls to ensure that it is isostatically supported; thus avoiding the possibility of seizure.

Modifications to the proposed control system to accommodate the I.P. pre-compaction cycle, are minor. The pump output would simply be piped directly to the HPA, thus by-passing the intensifier and accumulator. This may be conveniently achieved by using a standard solenoid operated 3-way valve. Allowance for increased flow should be made in the design of the HPA.

The decision as to whether a single press or two separate isostatic pressing systems will be employed will ultimately be made by the user. For the current design, attention is devoted to the H.P. control system, but provision is made where relevant, for simple conversion to a dual purpose system.

6.6 The High Pressure Chamber

The H.P. chamber is required to accommodate the pre-compacted pellet and to safely withstand the maximum working pressure.

The following constraints have been imposed by the user:-

(i) The 'bursting pressure' should not be less than 100,000 p.s.i.

(ii) The inside diameter and internal height should not be less than 8.24 in.

(iii) The outside diameter of the cylindrical wall (excluding the lower flange, Fig. 26) should not exceed 18in.

6.6-1 The chamber, material and fluid volumes

For a pre-compacted pellet nominally 8" in diameter and height, a chamber with nominal internal dimensions of 8.24 in. in diameter and height will allow an initial clearance* of 0.12 in. above and below the pellet, and all round the pellet in the radial direction. The volume of this chamber is V_c = 440 cu. in.

If the final pellet volume, V_m, is 390 cu. in., the maximum fluid volume in the chamber will be V_f = 50 cu. in.

6.6-2 Variation of chamber gain with chamber bulk modulus

The chamber gain, k_c is given by equation (56), which is reproduced below:-

$$k_c = \frac{\beta_c \beta_m \beta_f}{V_o \beta_f \beta_m + V_f \beta_m \beta_c + V_m \beta_f \beta_c} \tag{56}$$

If the bulk modulus of the chamber, β_c, is made very large, such that $\beta_c(V_f \beta_m + V_m \beta_f) >> V_o \beta_f \beta_m$, then the latter quantity may be neglected, and equation (56) reduces to:-

$$\left(k_c\right)_{max} = \frac{\beta_m \beta_f}{V_f \beta_m + V_m \beta_f} \tag{56-1}$$

Equation (56-1) represents the maximum value that k_c can theoretically attain for given values of the bulk moduli β_m and β_f and the assumed geometry. This relationship becomes exact when the bulk modulus of the chamber, β_c, becomes infinite, which is clearly not realizable in practice.

---

* Such clearance includes the plastic coating of the pellet which is of the order 0.003 in. thick.

If $\beta_f = 3 \times 10^5$,* $\beta_m = 3 \times 10^6$, $V_f = 50$, and $V_m = 390$; then from equation (56-1), $k_{c\ max} = 3400$.

A plot of the variation of the chamber gain k_c with the chamber bulk modulus appears in Fig. 27. This is the graphical representation of equation (56), with β_m, β_f, V_o, V_c, and V_f all constant and of magnitudes defined in the foregoing.

Examination of Fig. 27 reveals that as the chamber bulk modulus is increased from zero to about 6×10^6 p.s.i., the chamber gain, k_c, increases relatively rapidly. As β_c is further increased from 6×10^6 p.s.i. to 12×10^6 p.s.i., the chamber gain increases far less rapidly; (the increase in k_c as β_c changes from 6×10^6 p.s.i. to 12×10^6 p.s.i. is about 10% of the increase in k_c as β_c changes from 0 p.s.i. to 6×10^6 p.s.i.).

For values of β_c in excess of 12×10^6 p.s.i., k_c increases very slowly, thus doubling the bulk modulus from 12×10^6 to 24×10^6 p.s.i. increases k_c by only 4%.

It was shown in section 6.5, that for minimum power consumption, k_c should be as large as possible. The value of k_c will however be limited by chamber bulk moduli obtainable with practical vessel configurations.

An all steel monobloc chamber construction, and a steel-tungsten carbide duplex vessel configuration are examined in the following sections, and the chamber bulk modulus of each type is determined.

---

* Practical fluids with higher bulk moduli are not readily available.

6.6-3 Examination of a monobloc chamber construction

6.6-3-1 Diametral dimensions required for adequate strength:-

If the chamber is of a monobloc construction, then neglecting end effects the classical equations of Lamé for the stress distribution in a single thick-walled cylinder can be expressed as:-

$$\sigma_c = a + \frac{b}{r^2} \tag{57}$$

$$\sigma_r = a - \frac{b}{r^2} \tag{58}$$

where r = radius

σ_c = circumferential principal stress at radius r

σ_r = radial principal stress at radius r

a, b = constants.

The following additional terms are defined:-

p = internal chamber pressure

r_i = internal radius of cylinder bore

r_o = external cylinder radius

σ_{ci} = circumferential stress at the bore when pressure, p, is applied

σ_{ri} = radial stress at the bore when pressure, p, is applied

σ_{ro} = radial stress at outside wall of the vessel

f_t = uniaxial tensile strength of the material.

Inspection of equations (57) and (58) reveals that the principal stresses σ_c and σ_r reach their maximum tensile and compressive values respectively when r is minimum, i.e. at the bore. Failure may therefore be expected to occur at the bore.

The outside wall of the cylinder is a free surface, thus when

$r = r_o$, $\sigma_{ro} = 0$,

hence, from equation (58),

$$a = \frac{b}{r_o^2} \tag{58-1}$$

At the bore of the cylinder the radial stress is compressive and equal to the applied pressure.

i.e. when $r = r_i$, $\sigma_{ro} = - p$

hence, from equations (58) and (58-1),

$$p = b\left(\frac{r_o^2 - r_i^2}{r_o^2 r_i^2}\right) \tag{58-2}$$

From equations (57) and (58-1), the circumferential stress at the bore may be expressed as:-

$$\sigma_{ci} = b\left(\frac{r_o^2 + r_i^2}{r_o^2 + r_i^2}\right) \tag{57-1}$$

The constant, b, may be eliminated from equations (57-1) and (58-2) to yield:-

$$\frac{\sigma_{ci}}{p} = \frac{r_o^2 + r_i^2}{r_o^2 - r_i^2} \tag{59}$$

If failure at the bore occurs according to Tresca's criterion [22], then:-

$$\sigma_{ci} - \sigma_{ri} = f_t \tag{60}$$

and since $\sigma_{ri} = - p$, from equation (60), at failure,

$$\sigma_{ci} = f_t - p \tag{60-1}$$

Eliminating σ_{ci} from equations (59) and (60-1) and rearranging, the following relationship is obtained:-

$$r_o^2 = r_i^2 \frac{f_t}{f_t - 2p} \tag{61}$$

If r_o and r_i are fixed, then for a given value of p, the minimum tensile strength of the material is given by:-

$$f_t = \frac{2pr_o^2}{r_o^2 - r_i^2} \tag{61-1}$$

Thus for a bursting pressure, p = 100,000 p.s.i., r_o = 9 in., r_i = 4.12 in.;

$$f_t = 252{,}000 \text{ p.s.i.}$$

Tensile strengths of this order and greater are obtainable using steels such as the K9L, NMCV-D.T.D., 'Hecla' 174 and Hykro series of steels.

6.6-3-2 Determination of the bulk modulus:-

The expression for the bulk modulus of a cylinder (equation (5)) derived in section 4.4 is not directly applicable in this case since it neglects axial deformation of the cylinder. In this case, movement in the axial direction is not constrained by an external load and must be taken into account.

The following notation is used:-

p = internal applied pressure

r = internal radius of cylinder

h = internal height of cylinder

Δr = radial displacement when pressure, p, is applied

Δh = axial displacement when pressure, p, is applied

V_o = internal volume of the cylinder

ΔV_c = change in cylinder volume when pressure, p, is applied

D_o = outside diameter of cylinder

D_i = inside diameter of cylinder

E = Young's modulus for the material

ν = Poisson's ratio for the material

σ_a = axial stress

when an internal pressure is applied, the change in volume is given by:-

$$\Delta V_c = \pi r^2 h - \pi (r + \Delta r)^2 (h + \Delta h)$$

Expanding the above expression, and ignoring second order quantities,

$$\Delta V_c = \pi (2rh\Delta r + r^2 \Delta h) \tag{62}$$

The chamber bulk modulus β_c has been defined as $\beta_c = \dfrac{V_o p}{\Delta V_c}$

hence from equation (62), $\beta_c = \dfrac{prh}{2h\Delta r + r\Delta h}$

therefore $$\frac{1}{\beta_c} = \frac{2\Delta r}{pr} + \frac{\Delta h}{ph}$$

or $$\frac{1}{\beta_c} = \frac{1}{\beta_{cr}} + \frac{1}{\beta_{ca}} \tag{63}$$

where $\beta_{cr} = \dfrac{pr}{2\Delta r}$ is defined as the 'radial bulk modulus'

and $\beta_{ca} = \dfrac{ph}{\Delta h}$ is defined as the 'axial bulk modulus'.

Since $r = \dfrac{D_i}{2}$, β_{cr} is identically the expression obtained in equation (5), which is reproduced below using the current notation

$$\beta_{cr} = \frac{E}{2} \left(\frac{D_o^2 - D_i^2}{(1 + \nu)D_o^2 + (1 - \nu)D_i^2} \right) \tag{5-1}$$

The axial stress is assumed to be uniformly distributed over the cross-sectional area of the cylinder and is given by:-

$$\sigma_a = \frac{pD_i^2}{D_o^2 - D_i^2}$$

The axial strain $\frac{\Delta h}{h}$ is thus, $\frac{\Delta h}{h} = \frac{pD_i^2}{E(D_o^2 - D_i^2)}$

hence

$$\beta_{ca} = \frac{ph}{\Delta h} = E \frac{(D_o^2 - D_i^2)}{D_i^2} \tag{64}$$

For D_o = 18 in.; D_i = 8.24 in.; $E = 30 \times 10^6$; and $\nu = 0.3$, the values of β_{cr} and β_{ca} calculated from equations (5-1) and (64) are:-

$$\beta_{cr} = 8.2 \times 10^6 \text{ p.s.i.} \quad \text{and} \quad \beta_{ca} = 113 \times 10^6 \text{ p.s.i.}$$

The chamber bulk modulus calculated from equation (63) is:-

$$\beta_c \simeq 7.65 \times 10^6 \text{ p.s.i.}$$

With this value of β_c, the chamber gain, k_c, may be calculated from equation (56)(or read from Fig. 27) to be k_c = 2850.

It should be noted that with an all steel vessel, the maximum value that the bulk modulus can theoretically attain is:-

$$(\beta_c)_{max} = \frac{E}{2.6} = 11.5 \times 10^6 \text{ p.s.i.}$$

This is determined from equations (5-1), (64) and (63) by letting $D_o \to \infty$. Higher bulk moduli are obtainable with finite cylinders if a material with a higher modulus of elasticity is used. This is further discussed in the following section.

6.6-4 Examination of a duplex chamber construction

Materials with moduli of elasticity greater than that of steel tend to be brittle and have a low tensile strength. A chamber made for example from tungsten carbide to the user's dimensions would burst at pressures of less than 100,000 p.s.i. By making the chamber a duplex construction with a carbide inner liner in a steel outer component, a compact vessel with a low compliance (high bulk modulus) may be obtained.

6.6-4-1 Optimum diametral dimensions required for adequate strength

The following notation is adopted for the purposes of the subsequent analysis:-

E = Young's modulus

ν = Poisson's ratio

f_c = compressive strength of material in uniaxial test

f_t = tensile strength of material in uniaxial test

$\lambda = \frac{f_t}{f_c}$ = strength ratio

$F = \frac{f_{ti}}{f_{to}}$ = ratio of tensile strengths of inner to outer component cylinders

p = radial pressure applied to inner bore

p_2 = radial pressure at interface radius due to interference alone

Δp_2 = change in interface pressure when p is applied at the inner bore

r = radius

r_1 = inner bore radius

r_2 = interface radius

r_3 = outer radius

x = r_1/r_3 = overall radius ratio

y = r_2/r_3 = radius ratio of outer cylinder

δ = radial interference at interface

σ_r = radial Principal Stress

σ_c = circumferential Principal Stress.

N.B. The physical properties of the inner and outer cylinders are distinguished by the subscripts 'i' and 'o' respectively.

For the present case it will be assumed that the inner and outer component cylinders are made from tungsten carbide (with 10% cobalt) and EN 30 steel respectively, with the following properties:-

E_i = 90×10^6 p.s.i.; E_o = 30×10^6 p.s.i.

ν_i = 0.21; ν_o = 0.3

f_{ci} = 848,000 p.s.i.; f_{co} = 212,000 p.s.i.

f_{ti} = 106,000 p.s.i.; f_{to} = 212,000 p.s.i.

λ_i = 0.125; λ_o = 1.

$$F = \frac{f_{ti}}{f_{to}} = 0.5 .$$

Parsons and Cole [27] propose a method of optimising the parameters of short duplex cylinders based on the Mohr criterion of failure - a criterion for which the degrees of brittleness of the materials of construction can be taken into account.

The reliability of the Mohr criterion and the method used is confirmed by experimental work carried out by the authors. In all cases of 'semi-brittle' inserts in ductile outer cylinders slight discrepancies between predicted and actual failure were on the side of safety [27].

The Mohr criterion of failure may be expressed as follows:-

$$\sigma_c - \lambda\sigma_r = f_t \tag{64}$$

It should be noted that for purely ductile materials, i.e. $\lambda = 1$; the Mohr criterion assumes the form of the Tresca criterion of failure due to maximum shear stress, while for completely brittle materials, i.e. $\lambda = 0$; the Mohr criterion is identical to the Rankine criterion of failure due to maximum principal stress.

In deriving design relationships for the compound vessel, the authors [27] use the Lame equations (equations (57) and (58)) to determine expressions for the radial and circumferential stresses at the bore of each component cylinder; taking into account the separate effects of prestressing during assembly and the application of an internal pressure, p.

The Mohr criterion of failure (equation (64)) is applied to each case, and by assuming that failure occurs simultaneously in both components, an expression for the internal pressure that will cause failure is obtained.

For the case of a 'semi-brittle' insert (i.e. $0 < \lambda_i < 1$) in a fully ductile outer cylinder (i.e. $\lambda_o = 1$), the internal pressure that will cause simultaneous failure in both components is given [27] by:-

$$\frac{p}{f_{to}} = \frac{(y^2 - x^2)F + y^2(1 - y^2)}{(y^2 + x^2) + \lambda_i(y^2 - x^2)} \tag{65}$$

By differentiating equation (65) with respect to y, the authors show that for a given value of x, p has a maximum value when:-

$$y^2 = -\left(\frac{1 - \lambda_i}{1 + \lambda_i}\right) x^2 + x \sqrt{\left(\frac{1 - \lambda_i}{1 + \lambda_i}\right)^2 x^2 + \frac{2F + (1 - \lambda_i)}{1 + \lambda_i}} \qquad (66)$$

For the present case, if the vessel is to be designed to burst at twice the working pressure, then p = 100,000 p.s.i. By substituting for p, f_{to}, F and λ_i in equations (65) and (66), and solving simultaneously for x and y, the following values are obtained:-

$$x = 0.500$$

$$y = 0.691$$

Thus for r_1 = 4.120 in., r_2 = 5.690 in., and r_3 = 8.240 in.

An expression for the necessary radial interference at the interface is derived by the authors [27] and may be written thus:-

$$\delta = r_2 f_{to}(1 - y^2) \left[\frac{(1 - \nu_i)y^2 + (1 + \nu_i)x^2}{2E_i(y^2 - x^2)} + \frac{(1 - \nu_o)y^2 + (1 + \nu_o)}{2E_o(1 - y^2)}\right] - \frac{2x^2 r_2 p}{E_i(y^2 - x^2)} \qquad (67)$$

The value of p to be used here is again the bursting pressure, p = 100,000 p.s.i. . By substituting for x, y, r_2, f_{to} and the elastic constants, the necessary interference is found to be:-

$$\delta = 0.0069 \text{ in.}$$

The necessary diametral interference is thus $\delta_D = 2\delta = 0.0138$ in.

In order to ensure that the interface radius assumes a value of 5.690 in.

after assembly, the interference should theoretically be apportioned in the ratio:-

$$\delta_{Do}/\ \delta_{Di} = \frac{(1 - \nu_o)y^2 + (1 + \nu_o)}{E_i(y^2 - x^2)} \Bigg/ \frac{(1 - \nu_i)y^2 + (1 + \nu_i)x^2}{E_i(y^2 - x^2)}$$

from which it follows that δ_{Di} = 0.0019 in and δ_{Do} = 0.0119 in.

where δ_{Di} = initial diametral <u>oversize</u> of the <u>outer diameter</u> of the <u>inner</u> cylinder

and δ_{Do} = initial diametral <u>undersize</u> of the <u>inner diameter</u> of the <u>outer</u> cylinder.

If the interference is proportioned in a different way, the error incurred in the final interface radius will be very small, so that for practical purposes the interference may be arbitrarily divided as long as it totals 0.0138 in.

6.6-4-2 Determination of the bulk modulus:-

In order to determine the bulk modulus of the vessel, it is necessary to ascertain the change in interference pressure, Δp_2, when an internal pressure, p, is applied.

This is given by [27] as:-

$$\Delta P_2 \left[\frac{y^2(1 - \nu_i) + x^2(1 + \nu_i)}{E_i(y^2 - x^2)} + \frac{y^2(1 - \nu_o) + (1 + \nu_o)}{E_o(1 - y^2)}\right] = \frac{2x^2\ p}{E_i(y^2 - x^2)} \qquad (68)$$

Substituting for x, y and the elastic constants, the following expression is obtained:-

$$\Delta p_2 = 0.391\ p.$$

Hence if a pressure p is applied, the change in pressure, Δp, across the inner liner is given by:-

$$\Delta p = p - \Delta p_2$$

$$\text{i.e.} \quad \Delta p = 0.709p.$$

The radial bulk modulus, β_{cr} in this case is $\beta_{cr} = \dfrac{\Delta p r_1}{\Delta r_1}$

where Δr_1 = change in bore radius, thus using a development similar to that in the derivation of equation (5), β_{cr} for this case may be expressed as:-

$$\beta_{cr} = \frac{E_i}{1.418} \left[\frac{D_2{}^2 - D_1{}^2}{(1 + \nu_i)D_2{}^2 + (1 - \nu_i)D_1{}^2} \right] \qquad (5\text{-}2)$$

where D_1 = inner bore diameter = 8.240 in.

D_2 = interface diameter = 11.380 in.

By substituting for the elastic constants, and the diameters; in equation (5-2) the radial bulk modulus is found to be:-

$$\beta_{cr} = 18.5 \times 10^6 \quad .$$

If the duplex chamber assembly has the general configuration illustrated in Fig. 26, the liner will not be subjected to any direct stress in the axial direction. A longitudinal shear stress at the outer wall of the liner will however be transmitted by the interface friction, and this will tend to restrain axial deformation of outer cylinder. If this restraining effect is neglected, a conservative estimate of the axial bulk modulus may be obtained, by following a development similar to the derivation of equation (64).

In this case, the axial bulk modulus, β_{ca}, is given by:-

$$\beta_{ca} = \frac{E_i(D_3{}^2 - D_2{}^2)}{D_1{}^2} \tag{69}$$

where D_3 = outer diameter of the outer cylinder = 16.480 in.

Substituting for E_i, D_1, D_2, and D_3; in equation (69) the value of β_{ca} is found to be:-

$$\beta_{ca} = 62.6 \times 10^6 \text{ p.s.i.}$$

Substituting for β_{ca} and β_{cr} in equation (63), the following value of the chamber bulk modulus is obtained:-

$$\beta_c = 14.3 \times 10^6 \text{ p.s.i.}$$

Thus the bulk modulus of the duplex construction is almost twice that of the monobloc type. The value of k_c may be calculated from equation (56).

Thus k_c = 3050, which is about 7% higher than the value obtainable with the monobloc construction.

The maximum hydraulic horsepower required in each case may be predicted from equation (54). For a chamber pressure of 50,000 p.s.i., and a pressure rise rate of 3,000 p.s.i./sec., 7.96 hp is required for the monobloc chamber and 7.45 hp for the duplex. The saving would be about 0.5 hp or 6.5% if the duplex chamber is used.

The saving in power is considered to be too small to justifiy the extra cost of a duplex chamber, and the use of a monobloc pressure vessel is therefore recommended.

6.6-5 Further chamber considerations

In practice the chamber will probably be capable of withstanding pressures greater than the 'bursting pressure' for which it has been designed, since the chamber cap will tend to resist radial deflection and thus reduce circumferential stresses in the walls of the vessel. The end effects will also tend to increase the radial bulk modulus; hence the neglecting of these influences is on the side of safety.

6.6-5-1 Fatigue resistance

It is well established that the fatigue life of a plain walled cylinder subjected to repeated internal pressure is limited [43 -46]. When the pressures used are substantially lower than the yield strength of the material, the fatigue life is prolonged. Morrison et. al. [43] found that vessels for which the stress did not exceed 65,000 p.s.i. could have an extremely long fatigue life (termed 'infinite' by the authors). The cylinders used by the authors were made from nitrided 'Hykro' steel. If a monobloc chamber is used for the current application, the maximum shear stress is half the principal stress difference at the bore;

$$\text{i.e.} \quad \tau_{max} = \frac{\sigma_{ci} - \sigma_{ri}}{2} \tag{70}$$

where τ_{max} = maximum shear stress

σ_{ci} = circumferential stress at the bore

σ_{ri} = radial stress at the bore.

If the working pressure is p_c, and the inner and outer radii are r_i and r_o respectively, as in section 6.6-3, then from equations (59) and (70), putting $\sigma_{ri} = -p_c$ and rearranging, the following expression is obtained:-

$$\tau_{max} = \frac{p_c r_o^2}{r_o^2 - r_i^2} \tag{71}$$

From equation (71), for $r_i = 4.12$ in, $r_o = 9$ in. and $p_c = 50{,}000$ p.s.i.,

$$\tau_{max} = 63{,}000 \text{ p.s.i.}$$

Thus if the chamber is made from one of the recommended steels, and preferably from nitrided Hecla '174' or Hykro, a very long fatigue life should be obtainable.

6.6-5-2 Transfer function of the chamber.

For the subsequent analysis of the control system it will be assumed that the chamber gain has a value, $k_c = 2850$. It is important, however that the chamber gain be measured in practice before final adjustments to the control system are made. Experimental determination of the chamber gain would require the metering of the flow into the vessel and measuring the rate of pressure rise.

Assuming zero initial conditions, the transfer function of the chamber may deduced from equation (3), (derived in section 4.4).
If q = rate of flow into the chamber, p_c = chamber pressure, the transfer function may be expressed as:-

$$\frac{p_c}{q} = \frac{k_c}{S} \tag{72}$$

6.6-5-3 Referred bulk modulus

The effective bulk modulus is usually used as a measure of the elasticity of a system. While the cylinder or chamber gains, k_c, may always be directly compared, the effective bulk moduli are not directly comparable if the fluid volume is different from the total volume (as in the case of isostatic pressing). For purposes of comparison between different systems, the 'referred bulk modulus', i.e. the effective bulk modulus referred to the fluid volume, should be used.

The referred bulk modulus, β_r, is defined by the following expression:-

$$\beta_r = k_c V_f = \beta_e \frac{V_f}{V_o} \qquad (73)$$

where β_e = effective bulk modulus

V_o = total volume

V_f = fluid volume

For $k_c = 2850$, and $V_f = 50$ cu. in., $\beta_r = 142{,}500$ p.s.i.

6.6-5-4 Chamber for I.P. and H.P. operation

If the same chamber is to be used for both the I.P. and H.P. compaction cycles, the removable 'blank' (Fig. 26) should be made from a material with a high bulk modulus, e.g. tungsten carbide*. The chamber itself may be designed using the same procedure as in the foregoing. In order to reduce weight, the use of two or more component cylinders in the construction of the vessel should be considered.

The chamber gain is likely to vary considerably during the I.P. cycle, owing to the compressibility of the material being compacted. Even though the control system may be designed to be stable over the complete range of chamber gain variation, its performance will be less than optimal. Possible methods of compensating for large variations in the chamber gain are briefly considered in section 17.

---

* The true bulk modulus (i.e. under isostatic conditions) of tungsten carbide is about 50×10^6 p.s.i., or approximately twice that of steel.

6.7 High Pressure Generation

The pressure generation system consists of an H.P. intensifier driven by an I.P. pump; and an H.P. accumulator. These are briefly considered below, first individually and then collectively.

6.7-1 Operating principle of the pressure intensifier

A pressure intensifier is essentially a cylinder with at least two sections of differing diameter. Lower pressure fluid, introduced into the large end, causes the piston to move and eject a smaller quantity of higher-pressure fluid at the small end. In the absence of friction the intensification ratio would be equal to the ratio of the large to the small piston areas.

Intensifiers may be of single or multi cylinder construction, and each cylinder may be single or double acting*. A single cylinder, double acting intensifier is illustrated in Fig. 28(a).

The intensifier shown has an electrically operated reversing mechanism (see Fig. 28 (b)) although hydraulically actuated reversing mechanisms are also commonly used.

With the piston in the position shown in Fig. 28(a), the solenoid 'S' of the four way valve is de-energised and fluid from the I.P. pump enters port 'A'. I.P. fluid also enters H.P. chamber 'X' through check valve 1. Port 'B' is vented to the reservoir through the four-way valve; thus the piston moves from right to left, discharging H.P. fluid from chamber 'Y' to the accumulator -HPA line via check valve 3. When the piston reaches the end of its stroke, it engages limit switch L.S.2.

---

* A single acting cylinder implies one working stroke per cycle while double acting means two working strokes per cycle.

L.S.2 is spring loaded and normally open (N/O), it is thus temporarily closed by the piston. This action completes the circuit (Fig. 28(b)) which energises the relay coil. The relay is a double - pole type, both poles of which are normally open. One pole is used as a 'hold circuit' which keeps the relay energised after L.S.2 is re-opened when the piston retracts. The other pole of the relay energises the solenoid 'S' which actuates the four-way valve thus connecting port 'B' to the I.P. supply and venting port 'A'. At this stage the piston commences to move to the right thus discharging H.P. fluid through check valve 2. When the piston approaches the end of its stroke it contacts normally closed limit switch L.S.1 which breaks the relay circuit thus de-energising the solenoid and returning the valve to its original position. The cycle then repeats itself.

6.7-2 Operating principle of H.P. accumulator

Pressure accumulators are used to perform two major functions in hydraulic circuits, namely that of hydraulic power storage and pulsation damping. Hydraulic accumulators consist essentially of an externally loaded variable volume chamber. They are distinguished by the type of external loading used. The principal types of loading used are dead-weight loading, mechanical spring loading and compressed gas loading.

A dead-weight loaded accumulator is the simplest form of hydraulic energy storage vessel. It comprises a vertical cylinder fitted with a piston carrying a large mass or ballast. Although this type of accumulator has the unique advantage that the fluid pressure is kept constant irrespective of variations in the volume of fluid in chamber; it is not widely used since the physical space occupied by the mass or ballast, particularly at high pressures, is very large.

In the spring-loaded type of accumulator, the piston works against one or more compression or tension springs. This type of accumulator is usually used only in low pressure, low volume applications.

Accumulators which make use of compressed gas loading are the most compact and most widely used type, for all pressure ranges [29]. Compressed gas loading may be achieved in several different ways. The hydraulic fluid and compressed gas are usually separated by diaphragm, flexible bladder or bag or by means of a floating piston. For the H.P. range; the floating piston type is used since it is possible to make use of a difference in area, as in the case of an intensifier, to maintain gas pressures at reasonably safe levels.

A differential area, compressed gas loaded H.P. accumulator is illustrated in Fig. 29. The gas is contained in a flexible bag which is attached to the supporting chamber wall at one point only, i.e. at the precharging port. The gas bag is separated from the fluid chamber by a floating piston which may be solid or hollow. The object of the bag is to minimise the loss of compressed gas, so that the accumulator need be re-charged with gas very infrequently.

Nitrogen is usually used as the precharging gas, as a safety precaution against explosions, and the vessel is usually fitted with a "keroset" or equivalent type of plug which melts in the event of a fire and allows the gas to escape. The vessel supporting the bag is usually designed to withstand at least five times the maximum gas pressure.

Precharging is effected using a very low powered high pressure booster, which may be electrically or shop-air driven. These units are available in portable form, thus a single booster may be used to precharge a large number of accumulators. The accumulator is precharged through a series of check valves. (Usually at least two check valves are used to minimise subsequent loss of compressed gas.)

In spring-loaded and compressed gas loaded accumulators, the hydraulic-fluid pressure varies as a function of the volume of fluid being stored. In the case of a compressed gas accumulator, the variation in pressure can be made insignificant by making the gas volume very much greater than the hydraulic fluid volume. This however is uneconomical, and most accumulators are designed to restrict fluid pressure variation to between 20% and 25% of the minimum pressure.

6.7-3 The H.P. pump

Positive displacement pumps of the radial and axial plunger types are commercially available for operating pressures of up to 10,000 p.s.i. Variable and fixed displacement types are obtainable. For the current application, the type of pump used is immaterial, provided that it can meet the flow and pressure requirements of the system. The pump should therefore be selected on the grounds of proven reliability and low cost.

For the current application, a fixed displacement axial plunger pump is likely to prove most favourable. The size of pump to be used is dependent on the intensifier, accumulator and requirements of the control system. It is essential that the pump, intensifier and accumulator are chosen so as to form a compatible sub-system; they are thus examined collectively below:-

6.7-4 The selection of a pump - intensifier - accumulator combination

In selecting a set of components for the H.P. generating system, consideration must be given to the requirements of the control system, life and reliability of the components and cost.

Commercially available accumulators generally wear very little and thus have a relatively long life. They are usually very reliable, but rather expensive; the cost increasing with increasing storage capacity.

Axial plunger pumps are designed for continuous operation at full load, and their cost increases almost in direct proportion to their hydraulic power output, for units in excess of 10 hp output capacity. The cost per horse-power of output is generally higher for smaller capacity pumps.

Unlike pumps, intensifiers are generally intermittently used in hydraulic circuits. Commercially available types are thus designed with intermittent usage in mind, and if used continuously may consequently be expected to have a shorter life than a pump with the same power output. For double acting intensifiers of the type described in section 6.7-1, which are capable of producing pressures of up to 75,000 p.s.i., the cost of the unit per c.i.s. of output is approximately constant for units of 1 c.i.s. capacity and larger. Smaller capacity units of this type generally have a higher cost per c.i.s. of output at the same pressure level.

One possible approach is to have a large accumulator and a relatively low capacity intensifier and pump. The accumulator would have to store sufficient pressurised fluid for the duration of the cycle, and would be re-charged between production runs. The cost of a large accumulator could be offset by the savings due to the utilisation of a lower capacity pump and intensifier. This approach will however undermine attempts to minimise non-productive time in mass production processes. The duration of the cycle will also be limited by accumulator capacity.

Another approach is to use a very much smaller accumulator, just large enough to act as a pulsation damper, and to smooth out the intensifier delivery; coupled with an intensifier - pump combination that can readily meet the demands of the control system.

There is apparently no compromise between the two approaches that will not limit the duration of the cycle. In order to keep the system as flexible as possible, the second approach is adopted.

All the pressure generating components are available in standard sizes. An accumulator of the type described in section 6.7-2, with a capacity of one-half imperial pint (17.3 cu. in.) should be adequate for the current application. (With a chamber gain of 2850, the accumulator if fully discharged could raise the chamber pressure by nearly 50,000 p.s.i. without assistance from the intensifier.) A typical accumulator of this type and size would have the following properties:-

Pre-charging gas pressure	= 20,000 p.s.i.
Gas pressure when fully charged with hydraulic fluid	= 25,000 p.s.i.
Hydraulic fluid pressure when fully discharged	= 60,000 p.s.i.
Hydraulic fluid pressure when fully charged	= 75,000 p.s.i.

When the rate of pressure rise demanded by the programmer is 3,000 p.s.i./sec., from equation (51), the flow into the chamber should be 1.05 c.i.s.

As the fluid is supplied at a pressure higher than the chamber pressure, there will be a dilation of volume flow rate from the high pressure source to the chamber. The flow rate that the H.P. source will be required to deliver is related to the flow into the chamber by the following expression:-

$$q = q_s \left[1 + \frac{(p_s - p_c)}{\beta_r} \right] \qquad (74)$$

where q_s = delivery of H.P. source

q = flow into the chamber

p_s = H.P. supply pressure

p_c = chamber pressure

β_r = referred bulk modulus of the fluid as defined in section 6.6-5-3.

The maximum delivery required from the H.P. source will occur when p_s = 60,000 p.s.i., p_c = 50,000 p.s.i. and q = 1.05 c.i.s.

From equation (74), under these conditons q_s = 0.98 c.i.s. for β_r = 142,500 p.s.i.

When the supply pressure is p_s = 75,000 p.s.i.; p_c = 50,000 p.s.i. and q = 1.05 c.i.s., then for β_r = 142,500, from equation (74), the required delivery will be q_s = 0.89 c.i.s.

It should be noted that at a supply pressure of 60,000 p.s.i., and a delivery of 0.98 c.i.s., the hydraulic horse power available is 8.8 hp (from equation (53)) while at a supply pressure of 75,000 p.s.i. and a delivery of 0.89 c.i.s. the hydraulic horse power available is 10.1 hp. In either case, the chamber and its contents store potential energy at a rate of only about 8 hp (see section 6.6-4-2). Most of the balance of the power is used to accelerate the fluid from the H.P. source to the chamber, and is lost in turbulence in the chamber [47].

The H.P. source should thus be capable of delivering 0.89 c.i.s. at 75,000 p.s.i., to meet the requirements of the chamber. To this required delivery should be added an allowance for leakages in H.P.A., accumulator and elsewhere, therefore a fluid consumption of the order 1 c.i.s. at 75,000 p.s.i. should be allowed for.

A flow as high as 1 c.i.s. will probably only be required for short periods during the cycle, and an intensifier with half this capacity would probably suffice. If the cycle is very demanding however, a 0.5 c.i.s. intensifier might have to operate almost continuously under load in order to keep the accumulator charged. An intensifier of 1 c.i.s. output capacity would only have to operate for half the working time of the smaller unit, and its expected life will therefore be doubled. For a commercially available intensifier of the type described in section 6.7-1, a 1 c.i.s. output capacity unit costs only about 45% more than a 0.5 c.i.s. unit, and will therefore be more economical in the long run. (This reasoning does not apply to units of capacity greater than 1 c.i.s. unit. The increase in the expected life would just cover the additional capital outlay, but not the interest on the capital.)

It is thus recommended that an intensifier capable of delivering 1 c.i.s. of fluid at a pressure of 75,000 p.s.i. be employed. An intensifier of this size can meet the required system demand, the accumulator being used only to smooth out the intensifier output, and to supply additional power when necessary, for rapid error correction, or during the initial portion of the H.P. cycle, when the pellet is not fully compacted.

The hydraulic horse-power output of the intensifier will be a maximum of 11.4 hp. Allowing for an intensifier efficiency of 90%, the I.P. hydraulic power input required is 12.6 hp. At 10,000 p.s.i., from equation (53), the delivery of the pump should be 8.3 c.i.s. Allowing 2.5 c.i.s. for the servo-valve would require a pump of 10.8 c.i.s. capacity. The closest standard size is an 11.1 c.i.s. pump, which requires a 20 hp electric motor drive.

It was mentioned earlier that pumps with hydraulic horse power outputs in excess of 10 hp have an approximately constant cost per horse power, which is less than the cost per horse power for smaller pumps. As in the case of the intensifier, this has been taken advantage of, and the fact that the pump will not often operate at full capacity should prolong its life.

The pump flow should be directed by a sequence valve which guarantees flow to the servo-valve circuit. The intensifier circuit may be fed from the secondary side of the sequence valve, and can be unloaded by a separate unloading valve when the accumulator is fully charged. This arrangement is further discussed in section 6.17 .

6.8 The Pressure Transducers

High pressure can be measured by several different means, which may broadly be classified as:-

(i) Dead weight gauges

(ii) Elastic deformation gauges (Bourdon tubes, strain gauge cells, and bulk modulus cells [48].)

(iii) Change in electrical resistance of some material when subjected to pressure [18, 49].

(iv) Phase changes in a material subjected to pressure [18].

Sensors which are to be used for control purposes are generally required to couple high accuracy with rapid response, reliability, and insensitivity to influences such as temperature changes, magnetic and electrostatic fields. Only a few of the pressure measurement techniques are therefore generally utilised as transducers for control purposes.

One of the most successful types of high pressure transducers is the manganin cell, which depends on the change in resistance of a coil of wire which is exposed to high pressure [50]. Manganin alloy has a relatively large, positive, linear, resistance - pressure relationship and is relatively insensitive to temperature fluctuations. Manganin cells usually comprise two coils of wire, one of which is active, while the other is a compensating coil. The active coil may be directly exposed to the high pressure environment, provided a non polar fluid is used. The dormant coil is isolated from the high pressure fluid, but is placed in close proximity to the active coil, and is used to compensate for gross ambient temperature changes.

The resistance change is detected as an out-of balance current in a bridge circuit, which has an output amplifier with adjustable gain.

Although manganin gauges have been found to be linear and reproducible in pressure measurements to better than 0.1%, [18, 50], precision of this order is unusual in control applications, since adiabatic temperature changes within the material introduced by pressure fluctuations affect the resistivity of the wire slightly, even when ambient temperatures are steady [50].

Industrial type manganin pressure transducers are available for pressures of up to 400,000 p.s.i. A typical 0 - 50, 000 p.s.i. transducer has a maximum error of less than 1% of full scale and negligible time constant. Transducers of this type are relatively inexpensive, easily incorporated into hydraulic systems and require no maintenance.

Since the required accuracy of the process for the current application has not been specified, it will be assumed that an industrial manganin transducer of the type discussed in the foregoing will suffice.

A 0 - 50,000 p.s.i. transducer is required for T_1 (Fig 25), while a 0 - 75,000 p.s.i. rating is necessary for transducer T_2 (Fig 25). It will be assumed that the output amplifiers of both transducers have been adjusted to give an overall sensitivity of 1 volt per 5,000 p.s.i.

The transducer gain, k_t, is thus:-

$$k_t = \frac{1}{5000} = 0.0002 \text{ volts/p.s.i.}$$

6.9 The Hydraulic Power Amplifier and Servo-Valve

The hydraulic power amplifier (HPA) is envisaged as a spool type valve, which is driven by the servo-valve. It is intended that the output flow (to or from the chamber), under steady-state conditions, be proportional to, and of the same order as, the output flow from the servo-valve, which in turn is proportional to the system error. The purpose of the HPA is to raise the pressure, and hence the power level of the servo-valve output rather than amplify the flow.

A design study of the HPA is presented below.

6.9-1 Principle of the HPA

The HPA is schematically illustrated in Fig. 30. If the spool of the servo-valve second stage is moved from its central position to the right, in response to an amplified error signal from the servo-valve pre-amplifier, oil from the L.P. source will flow into L.P. chamber 1 of the HPA while fluid is discharged from L.P. chamber 2. The spool of the HPA is thus forced to move to the right, uncovering port 2. Fluid from the H.P. source is thus able to reach the H.P. chamber through ports 1 and 2. If the servo-valve spool is displaced to the left, the HPA spool will also move to the left, thus venting the H.P. chamber to the reservoir via ports 2 and 3.

Leakage flow past the right end spool land is drained to the reservoir via port 3, while leakage past the left end spool land reaches port 3 through the hollow main spool.

6.9-2 Design of the main orifice

If it is assumed that the flow through the main orifice (port 2, Fig. 30) is turbulent, the flow to the H.P. chamber may be represented by the 'orifice law', equation (75).

$$q_c = C_d \, a \sqrt{\frac{2}{\rho}(p_s - p_c)} \tag{75}$$

where q_c = flow to the H.P. chamber

C_d = coefficient of discharge

a = orifice area

ρ = fluid density

p_s = supply pressure

p_c = chamber pressure.

The flow, q_{c1}, from the H.P. chamber, to the reservoir is similarly given by equation (75-1), when the spool is moved to the left.

$$q_{c1} = C_d \, a \sqrt{\frac{2}{\rho} \, p_c} \tag{75-1}$$

It should be noted that equation (75-1) assumes zero reservoir pressure.

The basic assumption that the 'orifice law' applies is justified subsequently.

The 'orifice law' assumes incompressible flow, the coefficient of discharge being calculated using the fluid density downstream of the orifice. In the current application, the H.P. chamber fluid density should be used in conjunction with coefficients of discharge published in the literature.

The density of the fluid downstram of the orifice varies with chamber pressure according to the following law [33]:-

$$\rho = \rho_o(1 + \frac{p_c}{\beta_r}) \tag{76}$$

where ρ_o = density at atmospheric pressure = 0.78×10^{-4} lb sec^2/in^4

β_r = referred bulk modulus of the fluid = 142,500 p.s.i. (see section 6.6-5-3).

Since the flow through the orifice, q_c and q_{c1} (equations (75) and (75-1)), is dependent on the reciprocal of the square-root of the density, the estimated flow rate will be relatively insensitive to density variations.

It may be shown that the use of an 'average' value of density of 1×10^{-4} lb-sec^2/in^4 (corresponding to a chamber pressure of 40,000 p.s.i. by equation (76)) will lead to a maximum error of ± 5% in the estimated flow rate, as the chamber pressure is varied from 10,000 p.s.i. to 50,000 p.s.i.

Merritt [33] asserts that a value of 0.60 may be used for the discharge coefficient of all sharp edged orifices, provided that the flow is turbulent and the flow area immediately upstream of the orifice is much greater than the orifice area.

For $\rho = 1 \times 10^{-4}$ lb-sec^2/in^4, p_s = 60,000 p.s.i., p_c = 50,000 p.s.i, and C_d = 0.60, the area, a, required to allow a flow of q_c = 1 c.i.s. may be determined from equation (75) to be:- $a = 1.18 \times 10^{-4}$ sq. in.

In order to ensure linearity between flow and spool displacement, the orifice should be rectangular in shape.

Full peripheral port configuration will not be practical in the current application owing to the small value of the required area. The orifice should hence take the form of longitudinal slits, of which there should be at least two, diametrically oppositely placed to ensure pressure balancing.

If w is defined as the total width of the slits (in the circume-fential direction), then for a spool displacement, x, from the central position,

$$a = wx \quad . \tag{77}$$

A negative value of x implies that fluid flows from the H.P. chamber through the orifice to the reservoir, and is applicable to equation (75-1).

Equation (77) is valid only for a critically lapped valve (Fig. 31). Thus, for a critically lapped valve, the orifice equations (75) and (75-1) may be respectively written as:-

$$q_c = C_d \; w \; x \sqrt{\frac{2}{\rho}(p_s - p_c)} \tag{78}$$

which is valid for $x \geqslant 0$,

and

$$q_{c1} = C_d \; w \; x \sqrt{\frac{2}{\rho} p_c} \tag{78-1}$$

which is valid for $x \leqslant 0$.

The "flow gain" of the valve, k_f, is defined as:-

$$k_f \equiv \frac{\partial q_c}{\partial x} \qquad \text{for } x > 0 \tag{79}$$

and

$$k_f \equiv \frac{\partial q_{c1}}{\partial x} \qquad \text{for } x < 0 \; . \tag{79-1}$$

Thus

$$k_f = C_d \; w \sqrt{\frac{2(p_s - p_c)}{\rho}} \qquad \text{for } x > 0 \tag{80}$$

and

$$k_f = C_d \; w \sqrt{\frac{2 \; p_c}{\rho}} \qquad \text{for } x < 0 \; . \tag{80-1}$$

Thus far the flow gain has not been defined at $x = 0$, since equations (80) and (80-1) are identical only for the special case of $p_s = 2p_c$, which is a highly unlikely condition.

It is also apparent that the flow gain can vary considerably (by a factor of up to 2.55) as p_s and p_c take on extreme values. The loop gain of the control system is directly proportional to the flow gain of the HPA (see section 6.11) and since a steady loop gain is required for optimum performance (see section 6.14), it is desirable that variations in the flow gain of the HPA be compensated for. This is effected by the adaptive gain compensator (see section 6.10).

From equation (80), k_f is smallest when $p_s - p_c$ has its lowest value, which is 10,000 p.s.i. for the present case. From equation (80-1) k_f is a minimum when p_c is at its lowest value, which is also 10,000 p.s.i. for the current problem.

Thus $(k_f)_{min} = 100\ C_d\ w \sqrt{\frac{2}{\rho}} = 8480\ w$ for $C_d = 0.60$ and $\rho = 1.0 \times 10^{-4}$ lb-sec^2/in^4.

If the adaptive gain compensator is designed to multiply the loop gain of the control system by a factor of $\frac{100}{\sqrt{(p_s - p_c)}}$ for $x > 0$, and by a factor of $\frac{100}{\sqrt{p_c}}$ when $x < 0$, variations in p_s and p_c will be nullified. A 'compensated flow gain', K_f which is independent of pressure may thus be defined as:-

$$K_f = 8480\ w \tag{81}$$

Equation (81) is valid for all x, provided that the adaptive gain compensator multiplies the loop gain by the appropriate factor at $x = 0$ as well. If for example x approaches zero from a positive value, then the appropriate compensation factor is $\frac{100}{\sqrt{p_s - p_c}}$ for $x \geqslant 0$.

If on the other hand, x approaches zero from a negative value a compensation factor of $\frac{100}{\sqrt{p_c}}$ will be applicable for $x \leqslant 0$. It should be noted that the choice of the minimum value of k_f as the point of unit compensation is arbitrary, but convenient in this case.

The adaptive gain compensator is further discussed in section 6.10.

The flow-displacement characteristics of a valve is highly dependent upon the lap of the valve (Fig. 31). Underlapped valves display non-linear characteristics and suffer from excess leakage. The characteristics shown in Fig. 31(a) represent the ideal case of $(q_c)_o = |(q_{c_1})_o|$ when the spool land is centred. Since the flow gains associated with q_c and q_{c_1} cannot be compensated simultaneously, Fig. 31(a) is valid for the special case of $p_s = 2p_c$. For any other relationship between p_s and p_c, a net flow into or out of the chamber will exist for the land in the central position. The control system will therefore move the spool to an off-central position so as to minimise the error. Thus under steady state conditions, the spool of an underlapped valve will seldom be centrally positioned. As the control system behaves as a proportional controller under steady-state conditions (see section 6.11), a steady state error will exist. This coupled with the excess leakage and non-linear flow gain makes the underlapped configuration the least desirable of the three.

The overlapped valve case leads to 'dead zone' and hence loss of control in the overlapped region, and thus steady-state error.

The critically lapped valve represents the optimum condition, since it leads to linear flow-displacement characteristics. Null leakage*,

---

* Null leakage is not to be confused with the total H.P. leakage past the land into the reservoir chamber of the valve.

i.e. leakage into and out of the orifice through the annular clearance between the land and the valve bore will influence the characteristics in all three cases.

In the case of the critically lapped valve, the null leakage is not subject to flow-gain compensation, and will thus affect the flow-displacement characteristics in the null region, for all values of $p_c \neq \frac{p_s}{2}$. This will lead to steady-state errors as outlined in the case of the underlapped valve. The steady-state errors will however be small if the null leakage is small.

The null leakage may be minimised by ensuring that the clearance between the spool land and the bore is as small as possible; and for a given orifice area, by making the orifice slits as narrow as possible.*

Making the slit widths narrow will lead to a lowering of the flow-gain, but this is of minor importance, since the loop gain may easily be adjusted in the low power (electrical) portion of the control system.

The minimum value of the slit width is restricted by the practical limitation of accurately machining fine slits. A width of 0.005 in. for each slit is within the range of conventional spark erosion machining techniques. Thus if two slits are used, each 0.005 in. wide, then $w = 0.010$ in., hence for $a = 1.18 \times 10^{-4}$ sq. in., from equation (77), $x = 0.0118$ in. The minimum slit length required to allow the passage of 1 c.i.s. of fluid at a pressure difference of 10,000 p.s.i. is thus 0.0118 in. When the pressure difference is 65,000 p.s.i., the adaptive

---

* Since a pressure gradient exists across the length of the spool land, a longitudinal slit, which is as narrow as possible clearly represents the minimum leakage case, for a given radial clearance between the land and the bore.

gain compensator will adjust the gap to $x = 0.0046$ in. for a flow of 1 c.i.s.

The validity of the basic assumption that the 'orifice law' is applicable may now be examined.

The Reynolds number, Re, may be written as:-

$$Re = \frac{\rho q_c D_h}{\mu a} \tag{82}$$

where D_h = hydraulic diameter (in)

μ = dynamic viscosity (lb-sec/in^2)(reyns)

The hydraulic diameter, D_h, is defined as four times the ratio of the total flow area to the total wetted perimeter.

Thus

$$D_h = \frac{2a}{w + 2x} \quad . \tag{83}$$

Thus the Reynolds number may be re-written as:-

$$Re = \frac{2\rho q_c}{\mu(w + 2x)} \quad . \tag{84}$$

The dynamic viscosity, μ, depends upon the fluid used, and in general varies markedly with temperature, and to a lesser extent, with pressure. This is particularly true of hydraulic fluids which consist of pure hydrocarbons. Additives, which reduce variations of viscosity with temperature and pressure have been developed and many 'synthetic' hydraulic fluids have evolved [54].

The sensitivity of viscosity to temperature and pressure is usually graphically illustrated on charts available from the manufacturer of the fluid. Formulae derived from the charts which approximate the variation of viscosity with temperature and pressure over a certain range are often used.

The viscosity - pressure relationship at a fixed temperature may be expressed as follows [31]:-

$$\mu_p = \mu_{po} e^{\alpha_1 p} \tag{84-1}$$

where μ_p = dynamic viscosity at gauge pressure p

μ_{po} = dynamic viscosity at atomospheric pressure

and α_1 = pressure-coefficient of viscosity at constant temperature.

The pressure-coefficient of viscosity may vary from 2×10^{-5} to 2×10^{-4} depending on the fluid, the temperature and the pressure range.

The viscosity - temperature relationship is also often expressed in exponential form [33], although the simpler Herschel relationship may be expressed as follows:-

$$\mu_T = \mu_{To} \left(\frac{T_o}{T} \right)^{\alpha_2} \tag{84-2}$$

where μ_T = dynamic viscosity at temperature T°F and atmospheric pressure

μ_{To} = dynamic viscosity at temperature To°F and atmospheric pressure

α_2 = temperature-coefficient of viscosity.

The temperature-coefficient of viscosity may vary from 1.5 to 3.5 depending on the fluid, and the temperature range.

Variations of α_1 with temperature and α_2 with pressure may be regarded as secondary [55], and equations (84-1) and (84-2) may be combined to yield:-

$$\mu = \mu_o \left(\frac{T_o}{T} \right)^{\alpha_2} e^{\alpha_1 p} \tag{84-3}$$

where μ = dynamic viscosity at temperature T and gauge pressure p.

μ_o = dynamic viscosity at temperature T_o and atmospheric pressure

and α_1 and α_2 are averaged over the applicable temperature and pressure ranges.

It must be emphasised that equation (84-3) is not a rigorous relationship and is valid as an approximation over specified ranges of pressure and temperature. It is apparent that viscosity decreases with increasing temperature and increases with increasing pressure.

For the current application, the fluid is assumed to be a V.H.P. alcohol-polyether base type with $\alpha_1 = 2.5 \times 10^{-5}$ (p.s.i.)$^{-1}$ for $p \leqslant 75{,}000$ p.s.i.; $\alpha_2 = 1.8$ for $70°F \leqslant T \leqslant 150°F$ and $\mu_o = 1.7 \times 10^{-6}$ reyns at $T_o = 100°F$.

It should be noted that the viscosity may vary from about $2\mu_o$ to about $0.5\mu_o$ as the temperature is varied from 70°F to 150°F at atmospheric pressure. Furthermore, the viscosity can increase by a factor of up to 6.5 as the pressure is varied from atmospheric to 75,000 p.s.i. at 100°F.

For the purposes of determining the validity of the orifice law, the temperature of the fluid will be assumed to be constant at $T = T_o = 100°F$.

There is experimental evidence to indicate that the orifice law, equation (75) is applicable at low Reynolds numbers, provided that the coefficient of discharge, C_d, is replaced by C_{d_1} [33].

Where $$C_{d_1} = \xi\sqrt{Re} \tag{85}$$

The quantity ξ depends on orifice geometry and is called the laminar flow coefficient.

A critical Reynolds number is defined by replacing C_{d1} by the coefficient of contraction of orifices in the turbulent regime. For sharp edged orifices, the coefficient of contraction is 0.611 [33, 51].

Thus $$Re_c = \left(\frac{0.611}{\xi}\right)^2 \tag{85-1}$$

where Re_c = critical Reynolds number.

Hence for $Re \leqslant Re_c$, the flow through the orifice may be calculated from equation (75) using $C_d = C_{d_1}$. For values of $Re > Re_c$, the flow can be treated as turbulent and determined from equation (75) using $C_d = 0.60$.

For the case of $Re \leqslant Re_c$, from equations (75) and (85),

$$q_c{}^2 = \frac{2a^2\xi^2\ Re}{\rho}\ (p_s - p_c)\ . \tag{86}$$

From equation (84) and (86), substituting for Re,

$$q_c = \frac{4a^2\xi^2}{\mu(w + 2x)}\ (p_s - p_c)\ . \tag{87}$$

It should be noted that the flow is directly related to pressure difference, and since mass density is absent, dominated by fluid viscosity, as expected in laminar flow.

Wuest [52] has determined expressions for laminar flow through sharp-edged orifices.

For a rectangular slit (in this case the sum of two slits) of total width w and height x, where w >> x, the result is:-

$$q_c = \frac{\pi x^2 w}{32\mu}\ (p_s - p_c) \tag{88}$$

Thus from equations (87) and (88), by equating the flows the following expression is obtained:-

$$\xi^2 = \frac{\pi}{128}\left(1 + \frac{2x}{w}\right) \tag{89}$$

If term $\frac{2x}{w}$ is neglected, (this is consistent with the assumption that $w >> x$), then from equation (89), $\xi = 0.157$.

If $Re = Re_c$, then from equation (85-1), $Re_c = 15.2$.

The maximum absolute value of x at which this occurs may be determined from equations (75), (84) and (84-3); and (75-1), (84) and (84-3), by putting $C_d = C_{d1} = \xi\sqrt{Re_c} = 0.611$, $w = 0.01$ in., $\rho = 1.0 \times 10^{-4}$ lb-sec^2/in^4 and $\mu_o = 1.7 \times 10^{-6}$ reyns, then for $Re = Re_c = 15.2$,

$$|x| = 1.41 \times 10^{-3} \frac{e^{\alpha_1 p_s}}{(p_s - p_c)^{1/2}}$$

or $$|x| = 1.41 \times 10^{-3} \frac{e^{\alpha_1 p_c}}{p_c^{1/2}}$$ if equation (75-1) is used.

It may be shown that $|x|$ is maximum when $p_s = 60{,}000$ p.s.i., $p_c = 50{,}000$ at which $x = 6.3 \times 10^{-5}$in.

Since this small displacement represents the maximum value of x at which laminar flow can exist, the assumption that the orifice law, with $C_d = 0.60$ applies for all x will incur a negligible error. (The small value of x determined above also justifies the assumption that $w >> x$, which is presumed in equation (88)).

The fact that the orifice law is valid is indeed fortunate since the flow gain is dominated by density (which varies relatively slightly with termperature and pressure) as opposed to viscosity which varies significantly with temperature and pressure. Variations in a viscosity dominated flow gain would be difficult to compensate for with the result that optimum system performance may not be achievable.

6.9-3 Design of the spool

In order to ensure the validity of the orifice equations, so that the orifice is the controlling restriction, and to prevent flow saturation of the valve at less than maximum stroke, it is recommended, [33], that the flow passages have at least four times the maximum orifice area. If d is defined as the radial clearance between the bore and the stem joining the lands in the H.P. portion of the valve, then for two slits, of total width, w, and length x, the foregoing statement may be mathematically expressed as follows:-

$$4xd + wd \geqslant 4wx \; . \tag{90}$$

By making $d = w$, inequality (90) is satisfied for all x.

As the spool stem connecting the lands in the H.P. portion of the valve is hollow, it will behave as a thick walled cylinder subjected to external pressure. The H.P. acting on the exposed annular flanges of the lands will give rise to a tensile stress in the spool stem, which will be partially belanced by the resisting pressure on the spool ends. For a given spool bore diameter, the external diameter of the stem may be determined for adequate strength. Since the supply H.P. is likely to vary cyclically with a relatively high average frequency, the spool stem will be subject to fatigue. A factor of safety of at least three is recommended.

The radial and circumferential stress distributions in the stem may be determined using the Lamé equations (57) and (58) (see section 6.6-3-1).

By applying the boundary conditions, $\sigma_r = 0$ when $r = r_i$; and

$\sigma_r = -p_s$, when $r = r_o$; to equation (58), eliminating the constants a and b from (57) and (58) yields the following expressions:-

$$\sigma_r = -\frac{p_s r_o^2}{r_o^2 - r_i^2}\left(1 - \frac{r_i^2}{r^2}\right) \tag{91}$$

$$\sigma_c = -\frac{p_s r_o^2}{r_o^2 - r_i^2}\left(1 + \frac{r_i^2}{r^2}\right) \tag{92}$$

where σ_r = radial principal stress

σ_c = circumferential principal stress

p_s = supply pressure

r_o = radius of outside wall of stem

r_i = radius of inside wall of stem

r = radius $(r_i \leq r \leq r_o)$.

Both stresses are compressive for all radii, σ_c being the more compressive of the two at any particular radius. The maximum compressive stress, σ_{ci}, occurs at the bore and has a value given by:-

$$\sigma_{ci} = -\frac{2p_s r_o^2}{r_o^2 - r_i^2} \quad . \tag{92-1}$$

The axial stress, σ_z, is given by:-

$$\sigma_z = \frac{2\pi r_o d p_s - P_2 A}{\pi(r_o^2 - r_i^2)} \tag{93}$$

where d = radial clearance between stem and valve bore

A = spool end area

P_2 = resisting pressure in chamber 2 of the valve (Fig. 30) for $P_1 > P_2$.

The spool end area may be expressed as:-

$$A = \pi(r_o + d)^2 \ . \tag{94}$$

Hence substituting for A in equation (93), σ_z may be written as:-

$$\sigma_z = \frac{2r_o dp_s - (r_o + d)^2 p_2}{r_o^2 - r_i^2} \ . \tag{93-1}$$

If the Mohr criterion of failure is used (see section 6.6-4), it may be expressed as follows, using a factor of safety of 3.

$$\sigma_z - \lambda\sigma_{ci} = \frac{f_t}{3} \tag{95}$$

where f_t = tensile strength of the material in uniaxial test

f_c = compressive strength of the material in uniaxial test

λ = strength ratio = $\frac{f_t}{f_c}$.

It is apparent that a factor of safety of 3 is unattainable using any steel. This may be illustrated by assuming that $r_o >> r_i$. Thus from equation (92-1), the compressive stress σ_{ci} can never be less than twice the external pressure. If σ_z is neglected in equation (95), then substituting $\frac{f_t}{f_c}$ for λ, $f_c = - 3\sigma_{ci} = 6p_s$, i.e. f_c = 450,000 p.s.i. for p_s = 75,000 p.s.i. This would be the required compressive strength of an infinitely thick stem. For a stem with a practical geometry, a much higher compressive strength would be required. Steels with compressive strengths of this order are currently unavailable.

Compressive strengths of the required order are obtainable if tungsten carbide is used in the manufacture of the spool. It is recommended that this approach be adopted rather than compromising the factor of safety. The use of tungsten carbide in this application can also provide other benefits. These are discussed subsequently.

For given values of P_2, d, r_i, p_s, λ and f_t, it is possible to solve equations (92-1), (93-1) and (95) in order to establish r_o, the outer radius of the stem.

With reference to Fig. 30, the servo-valve acts as a pressure divider, so that:-

$$P_1 + P_2 = P_s \tag{96}$$

where P_1 = L.P. in chamber 1

P_2 = L.P. in chamber 2

and P_s = L.P. supply.

For $P_1 \geqslant P_2$, the load pressure drop, P_L, is defined as:-

$$P_L = P_1 - P_2 \quad . \tag{97}$$

Thus from equations (96) and (97),

$$P_2 = \frac{P_s - P_L}{2} \quad . \tag{97-1}$$

Hence for P_s constant, P_2 is smallest when P_L is maximum. If the servo-valve is chosen so that it operates only in its linear region (see section 4.6) (approximately 2/3 of the range) then the maximum load pressure will be approximately $(2/3)P_s$ [33].

Hence frome equation (97-1), for $P_L \simeq (2/3)P_s$, $P_2 \simeq (1/6)P_s \simeq 500$ p.s.i. for $P_s = 3{,}000$ p.s.i.

The radial clearance d has been fixed at d = w = 0.01 in. If r_i is chosen as 0.05 in. so as to allow free flow of leakage fluid to the reservoir, then for a spool made of tungsten carbide with f_t = 106,000 p.s.i. and λ = 0.125, and an H.P. supply of 75,000 p.s.i. from equations (92-1), (93-1) and (95)

$$r_o = 0.14 \text{ in.}$$

Substituting in equations (92-1) and (93-1), σ_{ci} = - 188,000 p.s.i. and σ_z = 13,400 p.s.i.

The circumferential stress at the bore is clearly the dominant stress.

The spool end area, A, may now be determined from equation (94). Thus for r_o = 0.14 in., d = 0.010 in.; A = 0.071 sq. in. Hence the nominal diameter of the spool end should be 0.3 in. This will also be the nominal diameter of the valve bore and the central land.

It has been shown in the foregoing, that by using tungsten carbide for the manufacture of the spool, a satisfactory factor of safety can be achieved with a practical spool geometry.

To ensure that seizure will not take place due differences in coefficients of thermal expansion, if close fits are used, the valve cylinder barrel should also be made of tungsten carbide, of a slightly different grade to ensure a good bearing surface. (The spool surface should preferably be slightly harder than the cylinder bore surface.)

Some of the other advantages of using tungsten carbide in the current application are:-

(i) Since the compressive stresses dominate, even for a lower factor of safety, it may be shown that a tungsten carbide spool will be lighter than a steel spool despite the fact that tungsten carbide is nearly twice as dense as steel. For example, if a unit factor of safety is chosen, a tungsten carbide spool would weigh about 25% less than a spool of the same length made from 'Hecla 174' steel. It will be shown that reduced weight can result in a higher natural frequency of the valve, which is desirable.

(ii) The coefficient of friction between the spool and the bore can be reduced by as much as 80% by using tungsten carbide rather than steel [33].

(iii) Owing to its high Young's modulus and low Poisson's ratio, tungsten carbide components deform far less than equivalent steel parts under the action of high pressure. For example, the axial strain in the stem, ε_z, is given [53] by:-

$$\varepsilon_z = \frac{1}{E}\left(\sigma_z - \nu(\sigma_c + \sigma_r)\right) \tag{98}$$

where E = Young's modulus

ν = Poisson's ratio.

For the current case, noting that $\sigma_c + \sigma_r = \sigma_{ci}$ (from equations (91), (92) and (92-)), for $E = 90 \times 10^6$ p.s.i., and $\nu = 0.2$; $\varepsilon_z \simeq 5.7 \times 10^{-4}$. For a steel stem of the same dimensions, it may be

shown that for $E = 30 \times 10^6$ p.s.i., and $\nu = 0.3$;
$\varepsilon_z \simeq 2.3 \times 10^{-3}$, an increase by a factor of about 4. Tungsten carbide is also less susceptible to creep.

(iv) Lands of smaller diameter are possible with tungsten carbide rather than steel spools. This will lead to less leakage.

(v) Tungsten carbide parts with their superior hardness and low coefficient of friction can be expected to have a higher resistance to wear and hence a longer life.

The main disadvantage of a tungsten carbide spool and cylinder barrel is the higher manufacturing costs associated with sintering and diamond machining. It is believed that this disadvantage is heavily outweighed by the advantages discussed in the foregoing.

With the land diameters determined by the foregoing, attention may be focussed on the length and clearance of the main (central) land. The leakage flow from the H.P. portion of the valve past the main land into the portion of the valve which is at reservoir pressure, will be termed the "main land leakage".

If main land leakage flow is assumed to be laminar, then for the spool centrally located, the leakage flow is given by equation (44), which is reproduced below:-

$$Q = \frac{\pi}{12} \frac{Dd^3\Delta p}{L\mu} \tag{44}$$

where Q = main land leakage flow

d = radial clearance between the land and the bore

Δp = pressure difference

L = length of the land

μ = dynamic viscosity of the fluid.

It is desirable to minimise the main land leakage for two reasons; firstly to reduce hydraulic power wastage and secondly, to restrict uncontrolled leakage to the high pressure chamber via the sides of the orifice slits.

Inspection of equation (44) reveals that the most effective means of minimising the leakage is to make the radial clearance, d, as small as possible. By carefully lapping the mating parts, a radial clearance of 0.0001 in. is obtainable. Such clearance will allow a smooth sliding fit between tungsten carbide components of nominal diameter 0.3 in. and land lengths of the order 0.25 in. Longer lands are not recommended, as they may inhibit the precision of the fit, and will add to the spool mass.

Even under isothermal conditions, the viscosity of the fluid will vary as the leakage fluid flows from the H.P. portion of the valve to region at reservoir pressure. This should be taken into consideration when estimating the main land leakage flow.

Equation (44) may be expressed in differential form as follows:-

$$Q\,\partial L = -\frac{\pi D d^3}{12\ \mu} \cdot \partial p \qquad (44\text{-}1)$$

where ∂p = Increase in pressure over length ∂L. (The negative sign in equation (44-1) implies that the pressure falls as the length is increased.)

From equation (84-3), for $T = T_o = 100°F$, $\mu = \mu_o e^{\alpha_1 p}$, where μ is the viscosity at pressure p.

Substituting for μ in equation (44-1) and integrating:-

$$Q \int_o^L \partial L = - \frac{\pi D d^3}{12\mu_o} \int_{p_s}^{o} e^{-\alpha_1 p} \partial p$$

Hence $$Q = \frac{\pi D d^3}{12 L \mu_o \alpha_1} (1 - e^{-\alpha_1 p_s}) \qquad (44\text{-}2)$$

For $D = 0.3$ in., $d = 0.0001$ in., $L = 0.25$ in., $\mu_o = 1.7 \times 10^{-6}$ reyns,

$\alpha_1 = 2.5 \times 10^{-5}(\text{p.s.i.})^{-1}$, and $p_s = 75{,}000$ p.s.i.,

from equation (44-2), $Q = 0.0063$ c.i.s.

If both H.P. lands have the same length and clearance, the total leakage would thus be 0.0126 c.i.s. or about 1.26% of the maximum intensifier output.

It must be emphasised that the leakage predicted in the foregoing can only be regarded as a rough approximation, as isothermal conditions are not likely to exist in the annulus.

It is not recommended that the clearance between the land and the bore be relaxed with the intention of reducing manufacturing cost at the expense of increased leakage. If the clearance is doubled, the leakage will be multiplied by a factor of eight, and under adverse temperature conditions, and wear in service, will approach intolerably high levels of leakage at a much earlier stage than necessary.

The life of the HPA will be governed by the leakage losses, and when these approach about 50% of the intensifier output, the intensifier may have to operate virtually continuously at full load in order to meet the requirements of a demanding pressure - time cycle. One way of extending the life of the HPA, would be to use a fluid of higher viscosity. This would however be at the expense of system performance since higher viscosities will lead to flow at lower Reynolds numbers through the orifice, with the result that the flow gain may be viscosity dominated even at large valve openings (see section 6.9-2).

Equation (44) and (44-2) are based on the assumption that the flow through the annular space between land and bore is laminar. This assumption may be justified by determining the Reynolds number from equation (82).

The hydraulic diameter, D_h, in this case is given by:-

$$D_h = \frac{4\pi Dd}{2\pi(D - d)} \simeq 2d \quad \text{for } D >> d.$$

The flow area in this case is $a_1 = \pi Dd$.

Substituting Q for q_c and a_1 for a in equation (82), and using the minimum value of μ, i.e. $\mu = \mu_o = 1.7 \times 10^{-6}$ reyns, the Reynolds number is found to be Re = 1.57 which undoubtedly indicates laminar flow.

It was also assumed that the spool is concentric with the bore in determining the leakage past the lands. If the annular channel between land and bore is eccentric, then equation (44) should be written as follows [33]:-

$$Q = \frac{Dd^3\Delta p}{12\mu L}\left(1 + 1.5\left(\frac{e}{d}\right)^2\right) \tag{44-3}$$

where e = distance between the bore centre and spool centre.

For the concentric case, i.e. e = 0, equation (44-3) is identical to equation (44). However for the extreme case of maximum eccentricity, i.e. e = d, the leakage will be increased by a factor of 2.5 over the concentric case.

Two factors can lead to eccentricity, firstly assymetrical slit placement (see section 6.9-4), and secondly lateral forces due to slight tapers on the lands. Taper on the land gives rise to non-linear pressure gradients along the length of the land which are sensitive to slight eccentricities. Blackburn [56], shows that the lateral forces are stabilising, i.e. they tend to centralise the spool if the small end of the taper is exposed to the higher pressure. On the other hand, the lateral forces are de-stabilising if the large end of the taper is exposed to the higher pressure and this tends to drive the spool against the bore wall. In the extreme case of relatively severe tapers, the phenomenon known as "hydraulic lock" can develop, in which the lateral forces are large enough to pin the spool to a wall despite the spool end forces [57, 58].

One method of ensuring concentricity and eliminating the possibility of hydraulic lock is to intentionally machine a taper to the higher pressure side of each land. This is considered to be impractical however, and will defeat the object of a lapped fit, or lead to an unnecessarily long and therefore heavy spool if additional tapered lands are used in tandem with existing 'parallel' lands.

The most commonly used method of achieving concentricity, involves the machining of annular grooves on the peripheral walls of the lands. Such grooves in effect 'short circuit' any localised pressure build-up by allowing flow around the spool periphery from high to low pressure areas [57, 58].

Such grooves are easily machined and should be used on all lands except the main one. Sweeney and Manham [57, 58] show that at least three balancing grooves should be used, while there is little advantage to using more than seven. It is recommended that the depth and width of the grooves should be at least ten times the radial clearance, and that the sides of the grooves should be perpendicular to the bore to prevent wedging of dirt particles.

Besides preventing hydraulic lock and reducing leakage by centring the spool, the balancing grooves also provide the following advantages:-

(i) The grooves act as reservoirs for dirt particles.

(ii) The entrapped fluid improves lubrication.

For the current application, it is recommended that seven balancing grooves be used on each of the lands (with the exception of the main land). The grooves should preferably be 0.003 in. wide and deep and equally spaced on a 0.031 in. pitch. (All lands should be 0.25 in. long and lapped to a radial clearance of 0.0001 in.).

The null leakage (i.e. the unintentional leakage into and out of the orifice through the annulus between land and bore when the spool is centralised); and the pressure sensitivity of the valve should be examined. As the radial clearance between the land and the bore exceeds the minimum displacement necessary for turbulent flow, the orifice law will govern the flow through the top (or bottom) of the slit.

The null leakage flow may be estimated from equation (78) and (78-1), if x is replaced by the radial clearance d.

From equations (78) and (78-1), the net null leakage, q_n, into the chamber is given by:-

$$q_n = q_c - q_{c1} = C_d w\ d \sqrt{\frac{2}{\rho}} \left[(p_s - p_c)^{1/2} - p_c^{1/2}\right] \qquad (99)$$

while the net null leakage, q_{n1}, out of the chamber is given by:-

$$q_{n1} = q_{c1} - q_c = C_d w\ d \sqrt{\frac{2}{\rho}} \left[p_c^{1/2} - (p_s - p_c)^{1/2}\right] \qquad (99\text{-}1)$$

From equation (99), q_n is maximum when p_s = 75,000 p.s.i. and p_c = 10,000 p.s.i. at which stage q_n = 0.0131 c.i.s.
From equation (99-1), q_{n1} is maximum when p_s = 60,000 p.s.i. and p_c = 50,000 p.s.i. for which it may be shown that q_{n1} = 0.0106 c.i.s. These null leakages are so small that they will impose negligible steady state errors (see section 6-11).

Furthermore, since the leakages through the ends of the slit exceed the main land leakage as previously determined, the null leakage through the sides of the slit will make a negligible contribution to the total null leakage, since the total slit width is a very small fraction of periphery of the bore. Thus, if the slit length is made equal to the land length, the null leakage will not be increased significantly. The valve would then be capable of handling very much higher flows and could suffice for I.P. operation if a dual-cycle system is used.

The magnitude of the "null pressure sensitivity" of a valve is a measure of the ability of the valve to clear away silt deposits which dam up the orifice when the spool is centred [33].

The "pressure sensitivity" is defined as $\psi_p \equiv \dfrac{k_f}{\psi_c}$ (100)

where k_f is the flow gain and ψ_c, the flow-pressure coefficient is defined as:-

$$\psi_c \equiv \frac{\partial q_c}{\partial (p_s - p_c)} \qquad \text{for flow into the chamber} \qquad (101)$$

and

$$\psi_c \equiv \frac{\partial q_{c_1}}{\partial (p_c)} \qquad \text{for flow out of the chamber.} \qquad (101\text{-}1)$$

The "null pressure sensitivity", ψ_{po} is the value of the pressure sensitivity when the spool is centred, i.e. at x = 0.

From equations (75) and (101),

$$\psi_c = \frac{0.5\ C_d\ w\ x\ \sqrt{2/\rho}}{\sqrt{p_s - p_c}} \qquad \text{for flow into the chamber} \qquad (102)$$

while from equations (75-1) and (101-1),

$$\psi_c = \frac{0.5\ C_d\ w\ x\ \sqrt{2/\rho}}{\sqrt{p_c}} \qquad \text{for flow out of the chamber.} \qquad (102\text{-}1)$$

From equations (80), (100) and (102),

$$\psi_p = \frac{2(p_s - p_c)}{x} \qquad \text{for flow into the chamber} \qquad (103)$$

while from equations (80-1), (100) and (102-1),

$$\psi_p = \frac{2p_c}{x} \qquad \text{for flow out of the chamber.} \qquad (103\text{-}1)$$

In either case, for a critically lapped valve, when x = 0, ψ_{po} should theoretically be infinite. In practice however, the null pressure sensitivity is always finite due to the radial clearance between the land and the bore. Thus x should be replaced by d for the null case. The minimum null pressure sensitivity will occur when $p_s - p_c = 10{,}000$ p.s.i. (with reference to equation (103)), or when $p_c = 10{,}000$ p.s.i. (with reference to equation (103-1)).

Thus $\left(\psi_{po}\right)_{min} = 2 \times 10^8$ p.s.i./in. for $d = 0.0001$ in.

This exceeds by two orders of magnitude, the minimum pressure sensitivity recommended by Merrit [33]. The valve should not therefore suffer from silting problems, unless the fluid is extremely dirty.

The design of the spool may be concluded by fixing the lengths of the stems which join the lands and deciding upon the outside diameter of the stem which operates in the reservoir pressure zone of the valve.

If the maximum stroke of the spool is set at 0.125 in. then stem lengths of 0.375 in.will allow annular H.P. entry and reservoir pressure exhaust ports 0.125 in. long. Ports of this size will permit free flow under all conditions. A valve stroke of 0.125 in. will allow the passage of 10.5 c.i.s. of fluid at a pressure difference of 10,000 p.s.i. The valve will therefore be able to meet ramp demands of at least 30,000 p.s.i./sec. without saturating. Furthermore if used for the I.P. cycle as well, it will be capable of satisfying ramp demands of the order 50 p.s.i./sec even if the chamber gain is as low as 10, for a 3,000 p.s.i. pressure difference across the orifice.

Since the stem which operates in the reservoir pressure zone is not subjected to severe pressures, its external diameter may be made considerably smaller than that of the other stem. An external diameter of 0.25 in. will allow a radial clearance of 0.025 in. between stem and bore, which will permit free flow to the reservoir.

It is recommended that the spool be drained by six radial holes, each 0.0625 in. in diameter drilled through the reservoir pressure stem in the positions shown on Fig. 32. The holes should be radially opposite for pressure balancing.

From equation (97-1), the maximum compressive load on the spool will occur when the load pressure, P_L, is zero, i.e. $P_2 = 0.5\ P_S = 1{,}500$ p.s.i. for a 3,000 p.s.i. supply. The maximum compressive stress in the reservoir pressure stem may easily be calculated by dividing the product of the spool end area and P_2 by the net stem cross sectional area (taking into account the drainage holes). It may be shown that the absolute maximum axial compressive stress is 5,000 p.s.i. The maximum axial deflection may be determined from the stress, Young's modulus, and length of the stem. The contraction in length may be shown to be of the order 2.1×10^{-5} in., or one order of magnitude less than the deflection in the H.P. stem.

The displacement transducer shaft should thus be attached to the land at chamber 2 (Fig. 30), i.e. nearest the reservoir pressure stem, so as to minimise the effect of changes in the spool length.

The recommended dimensions and configurations of the spool and sleeve are illustrated in Fig. 32. It should be noted that the lands have slightly different diameters to facilitate accurate lapping and simple removal of the spool. The orifice should be shaped as shown in order to ensure that it behaves as a sharp-edged orifice. It should be noted that the maximum overlap with the specified tolerances will not exceed the radial clearance, i.e. 0.0001 in.; this should not influence the valve behaviour.

The sleeve should be housed in an assembly of the form illustrated in Fig. 33. Such an assembly will ensure that the sleeve is not stressed by pressure differences across its walls, and that each different pressure is adequately isolated. It should be noted that all fluid entry and exit holes in the supporting cylinders are in the axial direction. Radial holes in thick walled cylinders subjected to a pressure difference across their walls, can give rise to tremendous stress concentrations, which may lead to premature fatigue failure [46].

The design of the thick walled supporting cylinders may be carried out using the methods discussed in sections 6.6-3 and 6.6-4. The design of the other housing details is straightforward and does not warrant further attention.

6.9-4 Flow forces on the spool

Referred to also as 'flow induced forces', 'Bernoulli forces', 'flow reaction forces' and 'hydraulic reaction forces', flow forces on the spool are as a result of fluid flowing in the valve chambers and through the orifice.

Two distinct flow forces may be identified as a steady-state flow force which is constant at a fixed valve opening and pressure difference, and a transient flow-force which exists only when the spool is in motion. The steady-state flow force will be examined first.

With reference to Fig. 34, the inherent fluid accelerating property of an orifice results in a jet force, F_j, acting normal to the plane of the fluid at the vena contracta. The jet force is given [33] by:-

$$F_j = \frac{\rho q_c^2}{C_c\ a} \tag{104}$$

where ρ = fluid density as previously defined

q_c = flow to the H.P. chamber

a = $w\ x$ = orifice area

C_c = coefficient of contraction for the orifice.

The following additional notation is adopted:-

a_1 = area of each of land faces (i) and (ii)

θ = jet angle

C_v = coefficient of velocity

C_d = coefficient of discharge

F_{j1} = steady-state flow force in axial direction

F_{j2} = steady-state flow force in radial direction.

The coefficients of discharge, velocity and contraction are connected by the law:-

$$C_d = C_v\ C_c \tag{105}$$

The equilibrant of the jet force F_j may be resolved into two perpendicular components such that:-

$$F_{j1} = -\ F_j \cos\theta \tag{106}$$

and

$$F_{j2} = -\ F_j \sin\theta\ . \tag{107}$$

The negative signs indicate that the reactions of the fluid flow forces are opposite to the general direction of flow.

The radial component of the force is balanced by the lateral component of the opposite slit. The axial force is not compensated and acts in a direction to close the valve port. This effect is illustrated by comparing the pressure distributions on land faces (i) and (ii) in Fig. 34.

By substituting equations (104), (75), (77) and (105) into equation (106), the following expression is obtained for the steady-state axial flow force:-

$$F_{j1} = - 2\ C_d\ C_v w\ x\ (p_s - p_c)\ \cos\theta \tag{108}$$

The value of $\cos\theta$ varies substantially with the orifice opening, and is dependent on the ratio of valve opening, x, to radial clearance, d.

The relationship between x, d and the jet angle, θ is given [33] by:-

$$\frac{x}{d} = \frac{1 + (\pi/2)\sin\theta - \log_e[\tan 0.5(\pi - \theta)]\cos\theta}{1 + (\pi/2)\cos\theta + \log_e[\tan 0.5(\pi/2 - \theta)\sin\theta} \tag{109}$$

The variation of θ and $\cos\theta$ determined from equation (109) is illustrated in Fig. 35.

It is apparent that the range of variation of $\cos\theta$ is from 0.933 when $x/d = 0$ to 0.358 when $x \gg d$. $\cos\theta$ falls very rapidly as x/d is initially increased, and has a value of about 0.44 when $x/d = 6$.

Equation (108) may be expressed as:-

$$F_{j1} = - k_j\ x \tag{108-1}$$

where $k_j = 2C_dC_v\ w(p_s - p_c)\cos\theta$. (108-2)

The steady-state flow force thus behaves as a centring spring, with a spring rate of k_j.

Since $(p_s - p_c)$ may vary from 65,000 p.s.i. to 10,000 p.s.i., for

$C_d = 0.60$, $C_v = 0.98$, $w = 0.01$ in., and taking into account variations in $\cos\theta$, the spring rate may vary between the extremes:-

$$42 \leqslant k_j \leqslant 710 \text{ lb}_f/\text{in.}$$

Owing to the tremendous variation in the flow force spring rate, very little attempt has been made in practice to exploit it as a centring spring. On the contrary, in cases where the spring rate is very large (for example when the port width, w, is very large), and small stroking forces are available, methods of flow force compensation are often applied; and mechanical springs or some other method is used to achieve centring action if required.

For the current design, steady-state flow force compensation will be unnecessary as the spring rate, k_j, is sufficiently small to be rendered negligible by the feedback action of the displacement transducer. This is verified subsequently.

The maximum stroking force required to overcome the steady-state flow-force, and achieve a flow of 1 c.i.s., may be determined from equations (75), (108) and (109). It may be shown to be 1.3 lb_f, which is negligible compared to the available stroking force, which with a load pressure of 2,000 p.s.i. and a spool end area of 0.071 sq. in. can be as high as 142 lb_f. Thus virtually all of the available stroking force will be used to accelerate the spool.

At small orifice openings, equation (108) should be written as follows [59];-

$$F_{j_1} = -2C_dC_v\, w\, (p_s - p_c)(\sqrt{x^2 + d^2})\cos\theta. \qquad (108\text{-}3)$$

The stroking force required to overcome the maximum null steady-state flow force may be determined from equation (108-3) by putting $x = 0$, when $(p_s - p_c) = 65{,}000$ p.s.i., and $\cos\theta = 0.933$.

The result is $F_{j_1} = -\ 0.0952\ lb_f$ which is negligibly small.

A similar analysis to that presented above may be used to establish the steady-state flow forces associated with flow through the main orifice from the H.P. chamber to the reservoir. The axial flow force in this case will clearly be in a direction to close the valve.

The transient flow force may now be examined. The column of fluid between the H.P. and main lands is accelerated when a change in spool position occurs. A force is therefore generated which reacts on the face of the spool lands. The direction of this force can be established by considering the acceleration of a small element of fluid in the direction of flow, as illustrated in Fig. 34. For an acceleration from left to right, the pressure on the left side of the element must be greater than that on the right side. Therefore the pressure on face (i) must be greater than that on face (ii). The transient flow force in this case acts to close the valve port; but this is not the general rule, since if flow is in the direction from the H.P. chamber to the reservoir, the transient flow force will tend to further open the valve port.

With reference to Fig. 34, if leakages are neglected and the fluid in the column is assumed to be incompressible, the transient flow force may be determined from Newtons second law as follows:-

$$F_t = -\ (\rho L_d a_1).\ \frac{d}{dt}\left(\frac{q_c}{a_1}\right) \qquad (110)$$

where F_t = transient flow force

L_d = distance between ports (Fig. 34)

$\rho L_d a_1$ = mass of fluid being accelerated

q_c/a_1 = instantaneous velocity of fluid in the axial direction.

Differentiating equation (75) with respect to time and substituting in equation (110), the following expression is obtained:-

$$F_t = - L_d C_d w\sqrt{2\rho}\left[\sqrt{p_s - p_c}\,\frac{dx}{dt} + \frac{x}{2\sqrt{p_s - p_c}}\,\frac{d}{dt}(p_s - p_c)\right]$$

or $F_t = F_{t_1} + F_{t_2}$ (111)

where $F_{t_1} = - L_d C_d w \sqrt{2\rho (p_s - p_c)}\;\frac{dx}{dt}$ (111-1)

and $F_{t_2} = - \frac{L_d C_d\, w\, x}{\sqrt{(2/\rho)(p_s - p_c)}} \cdot \frac{d}{dt}(p_s - p_c)$. (111-2)

If the supply pressure, p_s is constant, the term $\frac{d}{dt}(p_s - p_c)$ in equation (111-2) becomes $-\left(\frac{dp_c}{dt}\right)$.

The rate of chamber pressure rise will not ordinarily exceed 3,000 p.s.i., and since F_{t2} is maximum when $p_s - p_c$ = 10,000 p.s.i., $F_{t2\ max} = 3.18 \times 10^{-4}x$, for L_d = 0.25 in., C_d = 0.60, w = 0.01 in. and $\rho = 1 \times 10^{-4}$ lb_f-sec/in^4.

The factor 3.18×10^{-4} lb_f/in is a spring rate, and since $k_j \gg 3.18 \times 10^{-4}$; F_{t_2} may be neglected.

Equation (111-1) may be rewritten as:-

$$F_{t_1} = - B_1 \frac{dx}{dt} \qquad (111\text{-}3)$$

where $B_1 = L_d C_d w \sqrt{2\rho (p_s - p_c)}$. (111-4)

B_1 is a damping coefficient and is stabilising in the foregoing case as it opposes valve motion. In the reverse case, where flow is from the H.P. chamber to the reservoir, the coefficient of damping will be:-

$$B_1 = - L_d C_d w \sqrt{2\rho p_c} \tag{111-5}$$

In this case B_1 will be destabilising. The coefficient of damping can vary considerably with pressure and with changes in the damping length, L_d, as well as change sign, but since its value is very small (a maximum of 5.4×10^{-3} lb_f-sec./in.) it does not pose any serious difficulty. This is further discussed subsequently.

6.9-5 Frictional forces acting on the spool

As the spool is centralised by the balancing grooves, it does not contact the bore. The frictional force between the spool and the cylinder will therefore be viscous.

Dynamic viscosity is defined by the following expression:-

$$\tau \equiv \mu \frac{\partial u}{\partial y} \tag{112}$$

where τ = shearing stress in the fluid

u = velocity of fluid in x direction (i.e. $u = \frac{\partial x}{\partial t}$)

y = axis perpendicular to the x direction

$\frac{\partial u}{\partial y}$ = shear rate.

For a Newtonian fluid, the dynamic viscosity, μ, is independent of the shear. Most hydraulic fluids approximate Newtonian fluids [33]. In the narrow annulus of the current case, where the spool moves with a velocity $\frac{\partial x}{\partial t}$, the velocity profile is linear [47, 31], thus

$$\frac{\partial u}{\partial y} = \frac{1}{d} \frac{\partial x}{\partial t} \qquad \text{where } d = \text{radial clearance.}$$

And since $\tau = -\frac{F_f}{\pi DL}$, from equation (112),

$$F_{fi} = -\frac{\mu_i \pi DL}{d} \cdot \frac{\partial x}{\partial t} \tag{113}$$

where F_{fi}* = frictional force between the bore and the ith land

L = length of the land

D = diameter of the land

μ_i = dynamic viscosity of fluid adjacent to the ith land.

The dynamic viscosity of the fluid for the two inner lands is different from that for the two outer lands due to the pressure effect, while D, L and d are the same for all lands.

The total frictional force is given by:-

$$F_f = \sum_{i=1}^{4} F_{fi} = \frac{\pi DL}{d} \cdot \frac{\partial x}{\partial t} \sum_{i=1}^{4} \mu_i \tag{114}$$

The two inner lands are subjected to the same high pressure, while the two outer lands are substantially at comparable low pressures, the range of variation of which will not significantly alter the viscosity. The value of μ_i for the two outer L.P. lands may thus be approximated by μ_o.

Equation (114) may thus be written as:-

$$F_f = -\frac{2\pi DL}{d} \cdot \frac{\partial x}{\partial t} (\mu_o + \mu_a) \tag{115}$$

---

* The negative sign implies that the frictional force opposes the motion in the x direction.

where μ_a = average viscosity of the fluid adjacent to the H.P. lands.

For the land walls parallel to the bore, the pressure gradient is uniform, thus

$$\frac{dp}{dL} = - \frac{L}{p_s} \tag{116}$$

where p_s = H.P. supply pressure.

The average viscosity may be obtained from:-

$$\mu_a = \frac{1}{L} \int_o^L \mu_o \, e^{\alpha_1 p} \, dL \tag{117}$$

since $\mu_i = \mu_o e^{\alpha_1 p}$ at pressure p by equation (84-3), for $T = T_o$.

From equations (116) and (117),

$$\mu_a = - \int_{p_s}^{o} \frac{\mu_o}{p_s} \, e^{\alpha_1 p} \, dp \; .$$

Hence

$$\mu_a = \frac{\mu_o}{p_s \alpha_1} (e^{\alpha_1 p_s} - 1) \; . \tag{118}$$

The frictional force may thus be written as:-

$$F_f = - B_2 \frac{dx}{dt} \qquad (\text{since } \frac{dx}{dt} = \frac{\partial x}{\partial t}) \tag{119}$$

where $$B_2 = \frac{2\pi D L \mu_o}{p_s \alpha_1 d} (p_s \alpha_1 \, e^{\alpha_1 p_s} - p_s \alpha_1 + 1) \tag{120}$$

B_2 is the coefficient of viscous damping, and is always stabilising. For $D = 0.3$ in., $L = 0.25$ in., $\mu_o = 1.7 \times 10^{-6} lb_f\text{-sec./in}^2$, $p_s = 75{,}000$ p.s.i., $\alpha_1 = 2.5 \times 10^{-5}$ in^2/lb_f and $d = 0.0001$ in., a typical value of B_2 would be 0.044 lb_f-sec/in. B_2 represents a small damping coefficient, but is more than large enough in itself to ensure that any negative damping coefficients caused by the transient flow force is cancelled out.

6.9-6 Transfer function of the HPA

The equation of motion of the spool may be obtained by applying Newton's second law to the forces on the spool:-

$$P_L A + F_{j_1} + F_{t_1} + F_f = M \frac{d^2x}{dt^2} \tag{121}$$

The sum of the driving force, the flow forces and the friction force is equal to the inertia force.

Substituting equations (108-1), (111-3) and (119) into equation (121), the following expression is obtained:-

$$P_L A = M \frac{d^2x}{dt^2} + B \frac{dx}{dt} + k_j x \tag{122}$$

where M = mass of the spool and fluid being accelerated (including the fluid in the L.P. supply lines), $lb_f\text{-sec}^2/in.$

P_L = load pressure as defined in equation (97), p.s.i.

A = Spool end area (Fig. 30), sq. in.

$B = B_1 + B_2$ = net damping coefficient (B_1 is obtained from equation (111-4) and B_2 from equation (120)), lb_f-sec./in.

k_j = Flow force centring spring rate (defined by equation (108-2)), lb_f/in.

Assuming zero initial conditions, equation (122) may be Laplace transformed to yield:-

$$P_L A = (MS^2 + BS + k_j)x \; . \tag{123}$$

With reference to Fig. 30, applying the continuity equations to L.P. chambers 1 and 2 yields the following expressions:-

$$\bar{Q}_1 - cP_1 = \frac{dV_1}{dt} + \frac{V_1}{\beta_r}\frac{dP_1}{dt} \tag{124}$$

and

$$- cP_2 - Q_2 = \frac{dV_2}{dt} + \frac{V_2}{\beta_r}\frac{dP_2}{dt} \tag{125}$$

where Q_1 = flow from the servo-valve to L.P. chamber 1, c.i.s.

Q_2 = flow from the L.P. chamber 2 to the servo-valve, c.i.s.

V_1 = volume of L.P. chamber 1 and associated flow line, cu. in.

V_2 = volume of L.P. chamber 2 and associated flow line, cu. in.

P_1 = L.P. pressure in chamber 1, p.s.i.

P_2 = L.P. pressure in chamber 2, p.s.i.

β_r = bulk modulus referred to the fluid, p.s.i.

and c = leakage coefficient, (c.i.s.)/(p.s.i.) or in^5/lb_f-sec.

From equation (44),

$$c = \frac{\pi Dd^3}{12 L\mu_o} \tag{126}$$

μ_o is used for the viscosity, since at low pressures this will incur only a small error.

The volumes of the L.P. chambers may be written as:-

$$V_1 = V_{o1} + Ax \tag{127}$$

and

$$V_2 = V_{o2} - Ax \tag{128}$$

where V_{o1} = initial volume of L.P. chamber 1

and V_{o2} = initial volume of L.P. chamber 2.

If the spool is initially centred, then:-

$$V_{o1} = V_{o2} = \frac{V_t}{2} \tag{129}$$

where V_t = total volume of fluid in L.P. chambers and lines.

Differentiating equations (127) and (128), with respect to time yields

$$\frac{dV_1}{dt} = A\frac{dx}{dt} \tag{127-1}$$

and

$$\frac{dV_2}{dt} = -A\frac{dx}{dt} \tag{128-1}$$

Subtracting equation (125) from (124), and substituting for V_1, V_2, V_{o1}, V_{o2}, $\frac{dV_1}{dt}$ and $\frac{dV_2}{dt}$, the following expression is obtained:-

$$Q_1 + Q_2 - c(P_1 - P_2) = 2A\frac{dx}{dt} + \frac{Ax}{\beta_r}\frac{d}{dt}(P_1 + P_2) + \frac{V_t}{2\beta_r}\frac{d}{dt}(P_1 - P_2) \tag{130}$$

Since $P_1 + P_2 = P_s$ by equation (96), for the L.P. supply pressure, P_s, constant, $\frac{d}{dt}(P_1 + P_2) = 0$ (130-1)

The load flow, Q , is universally defined as:-

$$Q_L \equiv \frac{Q_1 + Q_2}{2} \tag{131}$$

Also, the load pressure, P_L, has been defined as $P_L = P_1 - P_2$ (97)

Substituting equations (130-1), (131) and (97) into equation (130) and Laplace transforming yields:-

$$Q_L = \frac{c}{2}P_L + \frac{V_t}{4\beta_r}P_L S + A\,x\,S \tag{132}$$

The load pressure, P_L, may be eliminated from equations (123) and (132) to yield the transfer function of the HPA.

Thus
$$\frac{x}{Q_L} = \frac{1/A}{S^3\left(\frac{V_t M}{4\beta_r A^2}\right) + S^2\left(\frac{V_t B}{4\beta_r A^2} + \frac{cM}{2A^2}\right) + S\left(\frac{cB}{2A^2} + \frac{V_t k_j}{4\beta_r A^2} + 1\right) + \frac{ck_j}{2A^2}} \qquad (133)$$

Allowing for about 4 in. of 3/32 in. nominal bore connecting pipe, the dead volume, V_t, of fluid in the L.P. chambers and lines will be of the order 0.046 cu. in.

Using a value of 1.4×10^{-3} $lb_f\text{-}sec^2/in^4$ for the density of tungsten carbide, and taking into consideration the mass of the fluid, and transducer rod, the total mass, M, will be of the order 1.6×10^{-4} $lb_f\text{-}sec^2/in.$

The leakage coefficeint, c, may be evaluated from equation (126), and will have a maximum value of 1.85×10^{-7} (c.i.s.)/(p.s.i.). The maximum value of the damping coefficient, B, as determined from equations (111-4) and (120) is 4.94×10^{-2} $lb_f\text{-}sec/in.$

The other data required is:-

$A = 0.071$ sq. in.

$\beta_r = 150{,}000$ p.s.i. (assumed)

$k_j = 710$ $lb_f/in.$ (maximum value, from equation (108-2)).

Using the data presented above, it may be shown that

$\frac{V_t B}{4\beta_r A^2} \simeq 7.6 \times 10^{-7}$ while $\frac{cM}{2A^2} \simeq 3 \times 10^{-9}$ $\quad \frac{V_t B}{4\beta_r A^2} >> \frac{cM}{2A^2}$ and the latter may be neglected.

Furthermore, $\frac{cB}{2A^2} + \frac{V_t k_j}{4\beta_r A^2} \simeq 0.01 << 1$.

Equation (133) may thus be written as:-

$$\frac{x}{Q_L} = \frac{1/A}{S^3\left(\frac{V_t M}{4\beta_r A^2}\right) + S^2\left(\frac{V_t B}{4\beta_r A^2}\right) + S + \frac{ck_j}{2A^2}} \qquad (133\text{-}1)$$

Finally, (133-1) may be expressed as:-

$$\frac{x}{Q_L} = \frac{1/A}{S\left(\frac{S^2}{\omega_2{}^2} + \frac{2\zeta_2 S}{\omega_2} + 1\right) + k_o} \qquad (133\text{-}2)$$

where $$\omega_2 \equiv \sqrt{\frac{4\beta_r A^2}{V_t M}} \qquad (134)$$

$$\zeta_2 \equiv \frac{B}{4A}\sqrt{\frac{V_t}{\beta_r M}} \qquad (135)$$

and $$k_o = \frac{ck_j}{2A^2} \qquad (136)$$

ω_2 is of the order 2.06×10^4 rads/sec.,

ζ_2 is of the order 0.077

while k_o varies in the range $7.6 \times 10^{-4} \leqslant k_o \leqslant 1.31 \times 10^{-2}$ rads/sec.

(This variation is due to changes in k_j only. Variations in c can cause k_o to take on values outside of this range.)

Since $\zeta_2 >> \frac{k_o}{2\omega_2}$ and $\frac{2\zeta_2 k_o}{\omega_2} << 1$, * the denominator of equation (133-2) may be approximately factorised in order to gain some insight into the dynamic behaviour of the HPA.

---

* The maximum value of $\frac{k_o}{2\omega_2}$ is $3.28 \times 10^{-7} << 0.077$. The maximum value of $\frac{2\zeta_2 k_o}{\omega_2}$ is $1.2 \times 10^{-7} << 1$. Hence the approximation is a very close one.

Thus

$$\frac{x}{Q_L} \simeq \frac{1/A}{(S + k_o)\left(\frac{S^2}{\omega_2^2} + \frac{2\zeta_2 S}{\omega_2} + 1\right)} \tag{133-3}$$

The various parameters in the transfer function may now be considered.

ω_2 is the hydraulic natural frequency caused by the interaction of the total mass and the spring rates of the two trapped fluid columns at the ends of the spool.

ζ_2 is the hydraulic damping ratio due to the viscous friction.

Owing to the fact that the spool mass and fluid dead volume are both very small while the referred bulk modulus and the spool end area are relatively large, the hydraulic natural frequency, ω_2, is extremely large. This is a favourable situation since the dynamic performance of the HPA will as a result be largely independent of the second order term, $\frac{S^2}{\omega_2^2} + \frac{2\zeta_2 S}{\omega_2} + 1$. Large variation in both ω_2 and ζ_2 may therefore be accommodated without sacrificing system performance. This is further discussed and illustrated in section 6.14.

The dynamic behaviour of the HPA will therefore be governed by the low frequency lag caused by the small value of k_o. This is undesirable for the following reasons:-

(i) In the extreme, if $k_o \simeq 0$, the lag term approximates a pure integration, and since the output from the HPA is again integrated by the cylinder, the open-loop transfer function will have a double integration in in the denominator and the system will be inherently unstable (see section 6.12). For small finite values of k_o, it will be possible to stabilise the system at very low loop gains; the transient response due to the large time constant, $1/k_o$, will however be extremely slow, with the result that it would not be possible to meet the requirements of the programmer.

(ii) The value of k_o is subject to such large variation that the performance of the system would be impossible to optimise.

There are several methods of overcoming the former difficulty: Increasing the width of the ports would increase the flow force spring rate, k_j; increasing the leakage coefficient, e.g. by shunting the servo-valve/HPA connection lines by means of a variable restriction which could be adjusted to raise k_o to a sufficiently high level. Finally, a phase lead compensator could be placed in series with the HPA, to effectively cancel the low frequency lag term and replace it by a higher frequency lag.

None of these methods will however eliminate the latter effect.

A widely used method of overcoming the former difficulty and significantly reducing the latter, is the loading of the spool with stiff mechanical springs. If such springs are arranged to act on the spool ends so as to oppose spool motion, they will effectively be in parallel with the flow force centring spring, and their spring rates will be additive.

If the mechanical spring rate is much larger than the flow force spring rate, the latter may be neglected, and a larger k_o, which is independent of k_j will be obtained. k_o will however still be subjected to variations in the leakage constant c. Other difficulties that will arise if mechanical springs are used are:-

(i) If the springs are incorporated within the HPA, the dead volumes of L.P. chambers will be increased, with a resultant lowering of the natural frequency ω_2.

(ii) If the springs are located external to the HPA, extension rods will have to be attached to the spool ends, and this will lead to additional sealing difficulties, further complications in the HPA design and increased spool mass which will again reduce the natural frequency ω_2.

(iii) It will be difficult to achieve an optimum system in practice, since many springs may have to be tried out in service, before an optimum combination is found. (Even though the optimum spring rate may be calculated, it is desirable in practice that the parameters which optimise the system performance be readily adjustable to allow for limitations and approximations in the mathematical model, and to facilitate changes in the optimisation criterion which may be required for specific applications.)

The utilisation of a displacement transducer to feedback the valve spool position to the servo-valve can overcome all difficulties associated with k_o, without imposing any of the disadvantages discussed in the foregoing. The displacement transducer is discussed below:-

6.9-7 The displacement transducer and the servo-valve

The displacement transducer is envisaged as a differential transformer type, with a moveable core. Commercially available transducers of this type usually have built in oscillators which are driven by an external constant voltage D.C. source. The output from the differential transformer is proportional to the displacement of the core from the central position. The output from the transformer is demodulated, filtered and fed to an output amplifier of adjustable gain.

For the current application it is recommended that a "Sanborn ± 0.020 in." type be used. A displacement transducer of this type can measure displacements of ± 0.020 in. from the central position with an error of less than 0.1%, and a displacement of up to ± 0.125 in. with an error of less than 1%; and negligible time constant.

As the servo-valve is used only to drive the HPA, a flow control device with a low delivery will suffice. The servo-valve should be reliable and have a rapid response. Two stage servo-valves such as the "American Brake Shoe Co., Aerospace Servo-Valve" have been found particularly suitable for low-flow applications of the present type [35]. Servo-valves of this type usually have a 'jet-pipe' first stage pre-amplifier which drives the main spool. Positional feedback is via a light centrally located leaf-spring.

For the current application a servo-valve of the type described in the foregoing with a rated flow of 2.5 c.i.s. and a flow-voltage gain, k_c, of 0.5 c.i.s./volt is considered suitable.

The transfer function of the servo-valve is given by:-

$$\frac{Q_L}{e_{v_1}} = \frac{k_e}{\frac{S^2}{\omega_1^2} + \frac{2\zeta_1 S}{\omega_1} + 1} \tag{137}$$

where Q_L = load flow

e_{v_1} = voltage input to the servo-valve

k_e = flow-voltage gain of the servo-valve = 0.5 c.i.s./volt

ω_1 = apparent natural frequency of the servo-valve

= 185 Hz. = 1160 rads/sec.

ζ_1 = apparent damping ratio = 0.7.

The servo-valve - HPA - displacement transducer combination constitutes the minor loop of the control system, and is represented as a block diagram in Fig. 36. The difference between the amplified system error, e_v, and the displacement transducer output, is amplified by the servo-valve input amplifier, which has a gain of k_i. The output from this amplifier, e_{v_1}, drives the servo-valve.

6.9-8 Transfer function of the minor loop

The transfer function of the minor loop is readily obtainable from the block diagram, Fig. 36, and may be written as follows:-

$$\frac{x}{e_v} = \frac{k_g}{S\left(\frac{S^2}{\omega_1^2} + \frac{2\zeta_1 S}{\omega_1} + 1\right)\left(\frac{S^2}{\omega_2^2} + \frac{2\zeta_2 S}{\omega_2} + 1\right) + \left(\frac{k_o S^2}{\omega_1^2} + \frac{2\zeta_1 k_o S}{\omega_1} + k_o\right) + k_g h} \tag{138}$$

where $k_g = \frac{k_i k_e}{A}$ (139)

k_g is the forward path gain of the minor loop and will have a value of 1000 if k_i is set at 142 for $k_e = 0.5$ and $A = 0.071$. h is the gain of the displacement transducer.

Equation (138) may be simplified by recognising that:-

$$\frac{2\zeta_1}{\omega_1} + \frac{2\zeta_2}{\omega_2} >> \frac{k_o}{\omega_1^2},$$

also $$1 >> \frac{2\zeta_1 k_o}{\omega_1}$$

and $k_g h >> k_o$. (h is of the order 0.5 - see section 6.13)

Hence all terms involving k_o may be neglected, and the minor loop becomes independent of this highly variable quantity. Equation (138) may thus be expressed as:-

$$\frac{x}{e_v} = \frac{k_g}{S\left(\frac{S^2}{\omega_1^2} + \frac{2\zeta_1 S}{\omega_1} + 1\right)\left(\frac{S^2}{\omega_2^2} + \frac{2\zeta_2 S}{\omega_2} + 1\right) + k_g h} \tag{140}$$

Equation (140) is the transfer function of the minor loop.

6.10 The Adaptive Gain Compensator

It was pointed out in section 6.9-2 that a means for compensating for variations in the HPA flow gain is required. It was shown that the gain compensator should multiply the loop gain by a function of the pressure drop across the orifice. The required gain factor is given by:-

$$Z = \frac{100}{\sqrt{\Delta p}} \tag{141}$$

where $\Delta p = p_s - p_c$ for $x > 0$ (141-1)

and $\Delta p = p_c$ for $x < 0$ (141-2)

where Z = compensation factor

p_s = H.P. supply pressure

p_c = H.P. chamber pressure.

The variation of the gain compensation factor, Z, with the orifice pressure drop, Δp, appears in Fig. 37 for the range $10{,}000 \leq \Delta p \leq 65{,}000$ p.s.i.

The curve may be conveniently subdivided into three regions, namely region III for $10{,}000 \leq \Delta p < 20{,}000$ p.s.i., region II for $20{,}000 \leq \Delta p < 36{,}000$ p.s.i. and region I for $36{,}000 \leq \Delta p \leq 65{,}000$ p.s.i.

A linear approximation of the curve may be obtained by replacing the arc of the curve in each region by a secant, as shown in Fig. 37. The linear approximation for Z is to within 3% of the exact value. This is sufficiently accurate for the current application (see section 6.14).

Each pressure transducer has a gain of $k_t = 0.0002$ volts/p.s.i. (see section 6.8).

Thus $V_s = k_t p_s$ (142)

$V_c = k_t p_c$ (142-1)

and $\Delta V = k_t \Delta p$ (142-2)

where V_s = Output of transducer T_2 (Fig. 25) from accumulator

V_c = Output of transducer T_1 (Fig. 25) from H.P. chamber

and $\Delta V = V_s - V_c$ for $x > 0$; or $\Delta V = V_c$ for $x < 0$.

The gain compensation factor is a decreasing function of the net transducer output, ΔV (Fig. 37). The circuitry of the function generator may be considerably simplified if Z is an increasing function of voltage.

Thus a synthetic voltage function, f(V), is defined, such that:-

$$f(V) = (28 - \Delta V) \tag{143}$$

The constant 28, represents the value of ΔV when the straight line approximation of region I on Fig. 37 is extrapolated to Z = 0.

The linearised variation of Z with f(V) is illustrated in Fig. 38. The function has a slope in regions I, II and III respectively of + 0.0258 for $15 \leqslant f(V) \leqslant 20.5$ volts; + 0.0530 for $20.5 < f(V) \leqslant 24$ volts; and + 0.176 for $24 < f(V) \leqslant 26$ volts.

The function generator is then a simple voltage dividing circuit as illustrated in Fig. 39. The transitions from region I to region II and from region II to region III is governed by the Zener diodes, ZEN 1 and ZEN 2 respectively. These do not conduct until their self-bias voltage has been exceeded.

The relationship between Z and f(V) is given by:-

$$\frac{Z}{f(V)} = \frac{R_o}{R + R_o} \tag{144}$$

where $R = R_1$ for region I, i.e. $\frac{Z}{f(V)} = 0.0258$;

$R = \frac{R_1R_2}{R_1 + R_2}$ for region II, i.e. $\frac{Z}{f(V)} = 0.0530$;

and $R = \frac{R_1R_2R_3}{R_1R_2 + R_2R_3 + R_3R_1}$ for region III, i.e. $\frac{Z}{f(V)} = 0.176$.

If R_1 is fixed at $1M\Omega$, then substituting the appropriate values of the ratio $\frac{Z}{f(V)}$, and the resistance R in equation (144) yields the following values for the resistances:-

$R_o = 26.5\ K\Omega$; $R_1 = 1M\Omega$; $R_2 = 838\ K\Omega$ and $R_3 = 170\ K\Omega$.

The required self-bias voltage of each Zener diode may be determined by evaluating the voltage dropped across the resistance R at each of the transition points.

i.e. Required Zener bias = $\frac{R}{R + R_o} \cdot f(V)$ at each transition point.

The required bias is 20.0 volts for ZEN 1 and 22.6 volts for ZEN 2. In practice, Zener diodes with the appropriate self-bias would have to be selected from a batch of 20 volt nominal self-bias.

The Zeners could be replaced if convenient by conventional diodes which are externally biased to the appropriate voltage by a potentiometer. The resistance of the used portions of the potentiometer should be taken into account when evaluating the resistance R_2 and R_3.

The output from the voltage divider is Z, the required compensation factor. The system error is multiplied by the factor Z using a high speed electronic multiplier, such as the type used in analogue computers [54]. Depending upon the manufacturer, multipliers usually divide the product of the two inputs by a factor of 50 or 100. This must be taken into consideration when assessing the forward path gain. If it is assumed that the multiplier in the current application has a reducing factor of 50, the output from the multiplier will be given by:-

$$e_o = \frac{Ze_i}{50} \tag{145}$$

where e_o = output from multiplier = modified system error

e_i = input to multiplier = system error

Z = input to compensation factor.

Electronic multipliers generally have negligible time constants and provide a true multiplication operation; i.e. the signs of the inputs are also algebraically multiplied [54].

The circuit for generating the function f(V) may now be examined. With reference to Fig. 39, B_1, B_2 and B_3 are electronic bi-stable elements. All three behave as normally open (N/O) switches until excited by an external signal. B_1 is closed if a small negative external signal is applied, and will remain closed even if this signal is removed. It may only be reset to its N/O position by a small positive external signal. B_2 and B_3 are similar in operation, but are closed by positive external signals and reset by negative external signals.

When x is positive, B_2 and B_3 conduct, and the voltage which appears at the adding amplifier is $(28 - V_s + V_c)$ volts which constitutes f(V). If x falls to zero B_2 and B_3 are reset. The function f(V) then becomes $(28 - V_c)$ volts, and is maintained until x becomes positive again.

6.11 Steady-State Response of the Control System

The major loop of the control system, Fig. 40, is constructed using equations (72), (81), (140) and (145).

Unity feedback may be established by ensuring that the input from the programmer is pk_t, where p = the instantaneous required pressure, and k_t = gain of the pressure transducer. The factor k_t, may then be incorporated in the forward path, thus yielding unity feedback. The transfer function of the system is readily obtainable from the block diagram, Fig. 40, and may be represented by:-

$$\frac{p_c(S)}{p(S)} = \frac{a_7}{a_1S^6 + a_2S^5 + a_3S^4 + a_4S^3 + a_5S^2 + a_6S + a_7} \tag{146}$$

where $a_1 = 1$ (146-1)

$a_2 = 2(\zeta_1\omega_1 + \zeta_2\omega_2)$ (146-2)

$a_3 = \omega_1{}^2 + \omega_2{}^2 + 4\zeta_1\zeta_2\omega_1\omega_2$ (146-3)

$a_4 = 2\omega_1\omega_2(\zeta_1\omega_2 + \zeta_2\omega_1)$ (146-4)

$a_5 = \omega_1{}^2\omega_2{}^2$ (146-5)

$a_6 = k_g h\omega_1{}^2\omega_2{}^2$ (146-6)

$a_7 = K\ k_g\omega_1{}^2\omega_2{}^2$ (146-7)

and $K = \dfrac{k_a\ k_t\ K_f\ k_c}{50}$ (146-8)

For $k_t = 0.0002$; $K_f = 84.8$; and $k_c = 2850$; from equation (146-8),

$$K = 0.966\ k_a \tag{146-9}$$

where k_a = gain of the system input differential amplifier.

6.11-1 Steady-State response to a step input

If the input to the system is a step function of magnitude p, such that $p(S) = \dfrac{p}{S}$, the steady-state output response, $p_c(t)_\infty$, is given by:-

$$p_c(t)_\infty = \lim_{S\to o} S \left[\frac{p\ a_7}{S(a_1S^6 + a_2S^5 + \ldots + a_6S + a_7)}\right]$$

hence $p_c(t)_\infty = p$ and since the error, $\varepsilon(S)$ of a unity feedback system is defined by:-

$$\varepsilon(S) = p(S) - p_c(S) \tag{147}$$

the steady-state error for a step input is zero.

Although the control system will not ordinarily be required to follow step inputs, the pressure time cycle may demand intervals of constant pressure. Thus the zero steady-state error to a step input property of the system is a highly desirable property. (In practice the steady state error can only be zero if the pressure transducer is 100% accurate. The actual steady state error will be approximately equal to the transducer percentage error.)

6.11-2 Steady-State error response to a ramp input

From equations (146) and (147) or directly from the block diagram, the error response of the system is given by:-

$$\frac{\varepsilon(S)}{p(S)} = \frac{a_1S^6 + a_2S^5 + a_3S^4 + a_4S^3 + a_5S^2 + a_6S}{a_1S^6 + a_2S^5 + a_3S^4 + a_4S^3 + a_5S^2 + a_6S + a_7} \tag{148}$$

If the input, p(S) is a ramp function of slope m, such that $p(S) = \frac{m}{S^2}$, the steady state error, $\varepsilon(t)_\infty$ obtained from equation (148) is given by:-

$$\varepsilon(t)_\infty = \frac{ma_6}{a_7} = \frac{mh}{K} \tag{149}$$

If a maximum error is specified, equation (149) could be used as a constraint on the ratio of h to K for the prupose of optimising the system. If for example the maximum allowed error is assumed to be equal to the error introduced by the pressure transducer, which may be 1%, then $\varepsilon(t)_\infty = 100$ p.s.i. when p = 10,000 p.s.i. For m = 3,000 p.s.i./sec, then from equation (149) K = 30 h. The constraint to be applied in such a case would therefore be:- $K \geqslant 30h$. A constraint of this order is very easily satisfied, in fact a more likely ratio of h/K would be of the order 1/200 (see section 6.13). Thus ignoring the error introduced by the transducer, steady state errors as low as 0.15% should be achievable.

6.11-3 Steady-State response to a flow disturbance

From the block diagram, Fig. 40, the output response, $p_c(S)$ to flow disturbance $q_d(S)$ in the position shown, is given by:-

$$\frac{p_c'(S)}{q_d(S)} = \frac{k_c(a_1S^5 + a_2S^4 + a_3S^3 + a_4S^2 + a_5S + a_6)}{a_1S^6 + a_2S^5 + a_3S^4 + a_4S^3 + a_5S^2 + a_6S + a_7} \qquad (150)$$

If the flow disturbance is a step of magnitude q_d, i.e. $q_d(S) = \frac{q_d}{S}$, the steady-state pressure response, $p_c'(t)_\infty$, ($p_c'(t)_\infty$ = contribution of q_d to the chamber pressure) is given by:-

$$p_c'(t)_\infty = \frac{q_d\ k_c\ a_6}{a_7} = \frac{q_d\ k_c\ h}{K} \qquad (151)$$

Using a value of 0.0131 c.i.s. for q_d (the null leakage flow, q_n, calculated in section 6.9-3), $p_c'(t)_\infty$ would impose an error of 1% on the chamber pressure (at 10,000 p.s.i. chamber pressure), if K = 0.374 h. For the case K/h = 200, the error would be of the order 2×10^{-4}% which is entirely negligible.

6.11-4 Steady-State response to a displacement disturbance

As any spool displacement disturbance occurs within the minor loop, the transfer function must be obtained from the block diagram of the minor loop, Fig. 36.

If the displacement disturbance is $x_d(S)$, then neglecting k_o,

$$\frac{x(S)}{x_d(S)} = \frac{S\left(\frac{S^2}{\omega_1^2} + \frac{2\zeta_1 S}{\omega_1} + 1\right)\left(\frac{S^2}{\omega_2^2} + \frac{2\zeta_2 S}{\omega_2} + 1\right)}{S\left(\frac{S^2}{\omega_1^2} + \frac{2\zeta_1 S}{\omega_1} + 1\right)\left(\frac{S^2}{\omega_2^2} + \frac{2\zeta_2 S}{\omega_2} + 1\right) + k_g h} \qquad (152)$$

The steady-state response to a step disturbance (e.g. due to stiction or changes in length of the H.P. portion of the spool) is therefore zero.

6.12 Absolute Stability Analysis

If the optimisation criterion is carefully chosen, the optimised control system will automatically be stable. Absolute stability information may however be applied with advantage to accelerate the optimisation procedure. The absolute stability criteria will thus be applied as constraints on the optimisation function.

The conditions for absolute stability may be established using the Routh-Hurwitz criterion [60]. A complete discussion of this criterion is beyond the scope of this study. The application of the criterion is however straightforward, and involves the ordering of the coefficients of the characteristic equation of the system,

$$a_1S^6 + a_2S^5 + a_3S^4 + a_4S^3 + a_5S^2 + a_6S + a_7 = 0 \qquad (153)$$

into an array as follows [61]:-

$$\begin{array}{c|cccc} S^6 & a_1 & a_3 & a_5 & a_7 \\ S^5 & a_2 & a_4 & a_6 & \\ S^4 & b_1 & b_2 & a_7 & \\ S^3 & c_1 & c_2 & & \\ S^2 & e_1 & a_7 & & \\ S^1 & f_1 & & & \\ S^0 & a_7 & & & \end{array}$$

where
$$b_1 = a_3 - \frac{a_1 a_4}{a_2} \qquad (154\text{-}1)$$

$$b_2 = a_5 - \frac{a_1 a_6}{a_2} \qquad (154\text{-}2)$$

$$c_1 = a_4 - \frac{a_2 b_2}{b_1} \qquad (154\text{-}3)$$

$$c_2 = a_6 - \frac{a_2 a_7}{b_1} \qquad (154\text{-}4)$$

$$e_1 = b_2 - \frac{b_1 c_2}{c_1} \qquad (154\text{-}5)$$

$$f_1 = c_2 - \frac{c_1 a_7}{e_1} \qquad (154\text{-}6)$$

By the Routh-Hurwitz criterion, a necessary and sufficient condition for stability is that there be no changes of sign in the elements in the first column of the array.

As a_1, a_2 and a_7 are all positive, stability will therefore be assured if b_1, c_1, e_1, and f_1 are all greater than zero.

It should be noted that if $h = 0$, then by equation (146-6), a_6 would be zero, so that c_2 (equation 154-4) would be negative. This would make f_1 negative (equation 154-6) and the system would be unstable. Thus the displacement feedback is essential for stability.

The stability criteria established in the foregoing are used to construct subroutine CONST (see Appendix A), which is used to restrict trial values in the optimisation programme to the feasible region.

6.13 Optimisation of the System

Inspection of equations (146), (146-6) and (146-7) shows that K governs the loop gain of the system, while h determines the system damping. These two parameters thus have the most significant effect on the system performance and the fact that they are readily adjustable makes them a convenient combination of variables for the purposes of optimising the system performance.

It should be emphasised at this stage that the object is not to achieve on-line optimal control, but rather to optimise the parameters of the system so as to obtain a good all round system performance.

6.13-1 The choice of an index of performance

The parameters of the system may be optimised according to an appropriate criterion, or index of performance.

The index of performance should take into consideration the particular cycle the system is required to follow. It could incorporate other factors, such as a weighted function of hydraulic power consumption; furthermore the performance index could be so weighted as to favour particular portions of the cycle, improving response and minimising errors in these regions at the expense of performance in other, less important areas.

A sophisticated performance index of the type described in the foregoing could only be synthesised if detailed information relating to the exact cycle; the importance of accuracy in the various portions of the cycle and the overall economics of the process is available. At present this is not the case, and the system will have to be optimised according to a more standard performance index. When the relevant information becomes available however, and a more appropriate performance index is consequently proposed it will be a relatively simple matter to adjust the parameters of the control system so as to satisfy the new criterion.

Several performance indices have been proposed, and have become generally accepted as standard quantitative measures of the important system specifications [39, 62, 63]. The index is applied to the response of a system to a step input, and gives rise to a positive number which is a measure of the performance. The best system is defined as that which minimises the performance index.

Some of the commonly used performance indices are defined as follows:-

$$\text{ISE} = \int_0^{T_1} \varepsilon^2(t)\ dt \tag{155}$$

$$\text{IAE} = \int_0^{T_1} |\varepsilon(t)|\ dt \tag{156}$$

$$\text{ITAE} = \int_0^{T_1} t|\varepsilon(t)|\ dt \tag{157}$$

$$\text{ITSE} = \int_0^{T_1} t\varepsilon^2(t)\ dt \tag{158}$$

where $\varepsilon(t)$ = system error at time t

T_1 = an arbitrary finite time chosen so that the integral approaches a steady-state value. This is often chosen as the settling time of the system.

ISE = integral of the square of the error

IAE = integral of the absolute magnitude of the error

ITAE = integral of time multiplied by the absolute error

ITSE = integral of time multiplied by the squared error.

The ISE and IAE criteria discriminate between systems which are excessively overdamped and underdamped, and the minimum value of each integral will occur at a different compromise value of the damping.

The ITAE and ITSE criteria reduce the contribution of the large initial error to the value of the performance integral while placing more emphasis on errors occurring later in the response.

Systems optimised according to the ITAE criterion generally give a favourable all round performance for a wide variety of inputs [35,63]. Furthermore, the ITAE criterion provides the best selectivity of the standard performance indices; i.e. it generally has a greater sensitivity to variations in the system parameters than the other forms [39]. A system optimised according to the ITAE criterion will therefore generally be closer to the optimum as determined by any of the other criteria, than would occur in the reverse case.

The ITAE performance index is adopted for the current problem, on the grounds of the foregoing arguments.

6.13-2 Representation of the control system in phase variable form

For purposes of minimising the ITAE index, so as to optimise the system performance, and also for the purpose of simulating the response of the system, it is necessary to express the output and error responses in the time domain. This may be accomplished by determining the inverse

Laplacé transform of equation (146) for a given input, p(S). As the form of the inverse transform is dependent upon the input, the simulation of the response to a sequence of different inputs, as in the case of a pressure-time cycle, will involve the determination of the inverse transform for each segment of the cycle. This approach is not amenable to a digital computer simulation of the transient response.

By expressing the overall transfer function of the system in phase variable form the tedious task of evaluating the inverse transform for a variety of inputs may be avoided. Furthermore, the state-vector differential equation may readily be expressed in a form suitable for solution on a digital computer. The simulation of the response to a sequence of different inputs may be relatively easily achieved, and the same basic computer programme may also be used in the optimisation of the system performance.

The vector-differential equation of the system may be established as follows:-

Equation (146) is reproduced below where $p_c(S)$ = output and p(S) = input:-

$$\frac{p_c(S)}{p(S)} = \frac{a_7}{a_1S^6 + a_2S^5 + a_3S^4 + a_4S^3 + a_5S^2 + a_6S + a_7} \qquad (146)$$

Cross multiplying this equation, and recognising that

$$p_c(S).S^n = \frac{d^n(p_c(t))}{dt^n} \equiv p_c^{(n)}(t)$$

the following expression is obtained:-

$$a_1 p_c^{(6)}(t) + a_2 p_c^{(5)}(t) + a_3 p_c^{(4)}(t) + a_4 p_c^{(3)}(t) + a_5 p_c^{(2)}(t)$$
$$+ a_6 p_c^{(1)}(t) + a_7 p_c(t) = a_7 p(t) \qquad (159)$$

The phase variables, X_1, X_2, ... X_6 are defined as follows:-

$$X_1 = p_c(t) = \text{output of the system} \qquad (159\text{-}1)$$
$$\dot{X}_1 = X_2 = p_c^{(1)}(t) \qquad (159\text{-}2)$$
$$\dot{X}_2 = X_3 = p_c^{(2)}(t) \qquad (159\text{-}3)$$
$$\dot{X}_3 = X_4 = p_c^{(3)}(t) \qquad (159\text{-}4)$$
$$\dot{X}_4 = X_5 = p_c^{(4)}(t) \qquad (159\text{-}5)$$
$$\dot{X}_5 = X_6 = p_c^{(5)}(t) \qquad (159\text{-}6)$$
$$\dot{X}_6 = p_c^{(6)}(t) = - a_2X_6 - a_3X_5 - a_4X_4 - a_5X_3 - a_6X_2 - a_7X_1 + a_7p(t) \qquad (159\text{-}7)$$

Equation (159-7) is derived from equation (159) and is valid as $a_1 = 1$ by equation (146-1).

The set of equations (159-2) to (159-7) may be expressed in vector-differential equation form as follows:-

$$\dot{X} = FX + Bp(t) \qquad (160)$$

where $\dot{X} = \begin{bmatrix} \dot{X}_1 \\ \dot{X}_2 \\ \dot{X}_3 \\ \dot{X}_4 \\ \dot{X}_5 \\ \dot{X}_6 \end{bmatrix}$; $X = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \\ X_5 \\ X_6 \end{bmatrix}$; $F = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ -a_7 & -a_6 & -a_5 & -a_4 & -a_3 & -a_2 \end{bmatrix}$; $B = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ a_7 \end{bmatrix}$

and $p(t)$ is a scalar quantity since there is a single input.

It should be noted that the column vector X constitutes a set of phase variables which are not the physical variables of the system. This is due to the fact that the vector differential equation was derived from the overall transfer function of the system, which is strictly an input-output relationship. Thus while X_1 represents the true system output, and p(t) the input, the phase variables X_2, X_3, ... , X_6 are not physically related to intermediate variables (voltages, displacement, flow, etc.) in the control system. The vector differential equation does however describe the overall input-output system behaviour.

The vector differential equation, equation (160) may be further simplified using a method suggested by Dorf [39]. The technique involves the introduction of a new column vector, M, such that the element M_1 generates the scalar input p(t). Using this method, the vector differential equation which describes the input may be derived.

Three examples are presented below, in order to illustrate the method:-

(i) For a unit step input

i.e. $p(t) = 1$

hence $M_1 = 1$

$\dot{M}_1 = 0$

The vector differential equation for M is thus:-

$$\dot{M} = [0]\, M \tag{161}$$

The initial condition is:-

$$M_1(0) = 1 \tag{161-1}$$

(ii) For a unit ramp input

i.e. $p(t) = t$

hence $M_1 = t$

$\dot{M}_1 = M_2 = 1$

$\dot{M}_2 = 0$

The vector differential equation for M in this case is given by:-

$$\dot{M} = \begin{bmatrix} \dot{M}_1 \\ \dot{M}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} M_1 \\ M_2 \end{bmatrix}$$

$$\text{or} \quad \dot{M} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} M \tag{162}$$

The initial conditions are:-

$$\left.\begin{aligned} M_1(0) &= 0 \\ M_2(0) &= 1 \end{aligned}\right\} \tag{162-1}$$

(iii) For a parabolic input

e.g. $p(t) = t^2 + t + 1$

hence $M_1 = t^2 + t + 1$

$\dot{M}_1 = M_2 = 2t + 1$

$\dot{M}_2 = M_3 = 2$

$\dot{M}_3 = 0$

The vector differential equation for M is thus:-

$$\dot{M} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} M \tag{163}$$

and the initial conditions are:-

$$\left.\begin{aligned} M_1(0) &= 1 \\ M_2(0) &= 1 \\ M_3(0) &= 2 \end{aligned}\right\} \tag{163-1}$$

It should be noted that in each case, the square matrix is governed only by the order of the polynomial, while the coefficients of the polynomial determine the initial conditions. Thus the square matrix corresponding to an nth order polynomial input will be an $(n + 1)*(n + 1)$ array with a unit superdiagonal and the other elements zero.

Harmonic inputs may be expressed in polynomial form using a Taylor's expansion.

The vector differential equations for X, and M may be combined into a single vector differential equation, by defining a new vector Y such that:-

$$Y = \begin{bmatrix} X \\ M \end{bmatrix} \tag{164}$$

and $$\dot{Y} = \begin{bmatrix} \dot{X} \\ \dot{M} \end{bmatrix} \tag{164-1}$$

Hence from equations (159-2) to (159-7) and (164-1)

$$\dot{Y}_1 = \dot{X}_1 = X_2$$

$$\dot{Y}_2 = \dot{X}_2 = X_3$$

$$\vdots \qquad \vdots \qquad \vdots$$

$$\dot{Y}_5 = \dot{X}_5 = X_6$$

$$\dot{Y}_6 = \dot{X}_6 = - a_2X_6 - a_3X_5 - \dots - a_7X_1 + a_7M_1 \quad (\text{since } M_1 = p(t))$$

$$\dot{Y}_7 = \dot{M}_1 = M_2$$

$$\dot{Y}_8 = \dot{M}_2 \qquad \text{(applicable for a ramp input function).}$$

The above set of differential equations may be expressed in vector differential equation form as follows:-

$$\dot{Y} = A\,Y \tag{165}$$

where A is a square matrix.

For a step input, the matrix, A is as follows:-

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ -a_7 & -a_6 & -a_5 & -a_4 & -a_3 & -a_2 & a_7 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{166}$$

While for a ramp input, the matrix, A is given by:-

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ -a_7 & -a_6 & -a_5 & -a_4 & -a_3 & -a_2 & a_7 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{167}$$

The solution to equation (165) is given by [39, 64]:-

$$Y(t) = \phi(t)\,Y(0) \tag{168}$$

where $$\phi(t) = e^{At} \tag{169}$$

$\phi(t)$ is the transition matrix.

The vector Y(0) represents the initial conditions.

For a unit step,

$$Y^T(0) = [\,0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 1\,] \tag{170}$$

while for a unit ramp,

$$Y^T(0) = [\,0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 1\,] \,. \tag{170-1}$$

The form of the solution to the new vector differential equation is that of an unforced system, which is driven only by the initial conditions.

In order to determine the transient response of the system, it is necessary to evaluate the transition matrix, $\phi(t)$.

6.13-3 Evaluation of the transition matrix

The matrix exponential function is defined as:-

$$e^{AT} = I + At + \frac{A^2t^2}{2!} + \frac{A^3t^3}{3!} + \ldots \tag{171}$$

where I = the identity matrix

A^2 = [A] [A], etc.

The series converges for any A, and for all finite t [65].

For digital computer application, it will be necessary to evaluate the phase variables, and transition matrix at specific time intervals, t = T. A truncated form of equation (171) will have to be used for digital computer application.

Thus $$e^{At} \simeq I + AT + \frac{(AT)^2}{2} + \ldots + \frac{(AT)^N}{N!} \tag{171-1}$$

where the series is truncated to N+1 terms.

The series will converge very rapidly if the interval T is chosen small enough so that $\lambda T < 1$, where λ = the absolute magnitude of the largest eigenvalue of the differential equation. It is not necessary to determine the eigenvalues in order to arrive at a suitable time interval. The sum of the eigenvalues is equal to the trace of the matrix F [65].

From equation (166), $\lambda \leqslant a_2$. If all the eigenvalues were equal, then $\lambda = \frac{a_2}{7}$. The applicable range of T is thus:-

$$\frac{1}{a_2} \leqslant T \leqslant \frac{7}{a_2}$$

The computer processing time increases as N increases and as T decreases. Furthermore, the number of terms required for convergence and thus N increase as T is increased. The interval T and number of terms N+1 should be chosen so as to assure a satisfactory convergence in the minimum amount of processing time. This is best done by trial and error. It is important that computer processing time be minimised in determining e^{AT}, since the transition matrix and the phase variables will be evaluated many times in the optimisation programme. This is further discussed subsequently.

The transition matrix and the phase variables are evaluated using subroutine OPTRAN (Appendix A), which also calculates the ITAE criterion. OPTRAN is also used for simulating the transient response of the system. Subroutine OPTRAN is described in Appendix A.

6.13-4 The method of optimisation

The parameters K and h of the system are optimised using a direct search technique (see Appendix B). The values of a_1, a_2, ..., a_7, used in subroutine CONST and subroutine OPTRAN are determined from equations (146-1) to (146-7), by subroutine INIT (Appendix A), using the values of ζ_1, ζ_2, ω_1, ω_2, k_g, K and h read into the main programme.

The values of K and h read in are the starting values. Subroutine CONST is then used to check that the starting values are in the feasible region. Thereafter the ITAE criterion is evaluated using subroutine OPTRAN. A search vector selects a new point in the K - h plane, and if this is found to be feasible by CONST, the ITAE criterion is again calculated and compared with the preceding value. If an improvement has occurred the direction of search is maintained and the search vector increased in length. Failure causes the search vector to change direction and decrease in length. The search vector ultimately homes in on the optimum point. The direct search technique is further discussed in Appendix B.

The following values were used for the system parameters:-

ω_1	= 1160 rads/sec.	(section 6.9-7)
ω_2	= 20,600 rads/sec	(section 6.6-6)
ζ_1	= 0.7	(section 6.9-7)
ζ_2	= 0.1	(section 6.9-6)
k_g	= 1000	(section 6.9-8)

Four sets of starting values were used for K and h, viz:-

(i) K = 160, h = 1

(ii) K = 160, h = 0.2

(iii) K = 60, h = 1

(iv) K = 60, h = 0.2.

On all occasions the optimum values reached were:-

$K = 112.3$; $h = 0.519$; the minimum value of the ITAE criterion being 1.457×10^{-5}.

For efficiency the programme was stopped during these runs when the step size had been reduced to 1% of the initial range. In order to determine the optimum more precisely, the programme was run once more using as a starting point the values of K and h determined in the optimisation above; an initial step length of 1% of the initial range, and a minimum step length of 0.05% of the range. The improvement was of the order 0.1% in the value of the ITAE criterion, and the values of K and h changed by less than 0.05%.

The fact that the optimum has been reached is finally confirmed by inspecting the sensitivity analysis (section 6.14).

For all practical purposes the optimum values of K and h may be taken to be 112.3 and 0.52 respectively.

Hence from equation (146-9), the differential amplifier setting should be $k_a = 116$.

Thus all the parameters in the system are defined.

6.14 Sensitivity Analysis

The sensitivity of the system performance to variations in some of the parameters is established using subroutine OPTRAN to calculate the value of the ITAE criterion (denoted CRIT in the programme) as the parameters are varied one at a time. This is achieved by placing the subroutine OPTRAN in a DO loop which varies the particular parameter under consideration while the others are held constant.

6.14-1 Sensitivity to variations in the gain, K

The variation of the measure of performance, CRIT with K is shown on Fig. 41 which is plotted for h = 0.519 and the other parameters as defined in section 6.13-4. The shape of the curve confirms the fact that the optimum occurs in the region of K = 112.3 . It is apparent that the system performance is relatively insensitive to small changes in the gain near the optimum point. Variations of the order ±5% in K will cause a loss in performance of the order 3%. Larger variations in K will cause a significant loss in performance.

6.14-2 Sensitivity to variations in the displacement feedback, h

The variation of the ITAE criterion with h is plotted in Fig. 42 for K = 112.3. Theshape of the curve confirms that the optimum value of h is of the order 0.52. Large deviations from the optimum point will result in significant loss in performance. As the displacement transducer is accurate to 0.1%, and may be precisely set operation near the optimum point is easily effected and optimum performance is approachable.

6.14-3 Sensitivity of the performance to variations in the natural frequency and damping ratio of the HPA

Inspection of Figs. 43 and 44 reveals that the performance of the control system is highly insensitive even to very large variations in the damping ratio ζ_2 and the natural frequency ω_2 of the HPA. This is desirable since both of these parameters may differ considerably from the calculated values. Furthermore, since the damping ratio depends on fluid viscosity, it may vary significantly during the course of a cycle.

6.15 Simulation of the Transient Response

6.15-1 Response to a unit step input

The transient response of the optimised system to a unit step input is determined using subroutine OPTRAN, with the square matrix, A defined by equation (166), and initial conditions defined by equation (170).

The state of the system is evaluated at intervals of 0.1 ms, and 18 terms are used in the calculation of the transition matrix (equation (171-1)). The programme prints the value of the output, $Y_1 = p_c(T)$, after each increment in time. (See Appendix A.)

The simulated output response is plotted in Fig. 45. The response is rapid, the 10% to 90% rise time being 4.4 ms. The settling time is of the order 20 ms while the percentage overshoot is approximately 6%.

6.15-2 Response to a 'plateau' input

In order to illustrate the method of simulating the response of the system to various input cycles, and at the same time, to determine the response to a ramp input, a 'plateau' input cycle has been chosen. The cycle consists of a ramp of slope +100 p.s.i./sec for $0 \leqslant t \leqslant 100$ ms; a region of constant input pressure, p = 10 p.s.i. for $100 < t \leqslant 200$ ms and a ramp of slope -100 p.s.i./sec for $200 < t \leqslant 300$ ms. A cycle of short duration was chosen so that the salient features of the response could be illustrated without unnecessary wastage of computer processing time. The choice of ramp slopes is arbitrary; steeper slopes accompanied by a higher plateau will merely alter the scale of the simulation rather than its form.

The simulation is carried out using programme 'TRANSIM' (see Appendix A) which applies subroutine OPTRAN successively to the three stages of the cycle.

For the first stage, the square matrix A, as defined in equation (167) for a ramp input is applied with initial conditions as follows:-

$$Y^T(0) = [\ 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 100\ ].$$

For the second stage, A is redefined as the matrix for a step input (equation (166)), and the state of Y at t = 100 ms = 0.1 sec. is used as the initial conditions; with one exception: the initial value of input generator, $M_1(0.1)$ is defined as 10 p.s.i.

The initial conditions for the second stage are thus:-

$$Y^T(0.1) = [\ Y_1 \quad Y_2 \quad Y_3 \quad Y_4 \quad Y_5 \quad Y_6 \quad 10\ ]$$

where Y_1, Y_2, ... Y_6 define the state at t = 0.1 sec.

For the third stage, A is redefined as equation (166), and the initial conditions this time are obtained from the values of the phase variables when t = 0.2 sec. The initial value of the input slope is $M_2(0.2) = -\ 100$ p.s.i./sec.

Thus $Y^T(0.2) = [\ Y_1 \quad Y_2 \quad Y_3 \quad Y_4 \quad Y_5 \quad Y_6 \quad Y_7 \quad -100\ ]$

where Y_1, Y_2, ... Y_7 define the state at t = 0.2 sec.

The programme is arranged to print the input and output every milli-second.

The results are plotted in Fig. 46. During the initial ramp the output approaches the input slope very rapidly and effectively attains the same slope after about 10 ms. The output continues to follow the input with a constant time lag of about 6 ms which results in a 'steady state' error of 0.46 p.s.i. When the input changes to a 'plateau' the output coincides with it after a short transient which results in a negligibly small overshoot. When the input begins to decrease during the final stage of the cycle the output slope again follows that of the input after a short transient which results in a time lag of about 6 ms.

The simulation of more complex cycles may be effected using the technique presented in the foregoing, and modifying programme TRANSIM accordingly.

6.16 Relative Stability of the Optimised Control System

A measure of the relative stability may be obtained by examining the open-loop frequency response of the optimised system. The open-loop frequency response may be simulated on a digital computer using programme 'FREQRES' (see Appendix C) which is capable of handling most standard transfer function forms.

From the block diagram, Fig. 40, the open-loop transfer function GH(S), is given by;-

$$GH(S) = \frac{a_7}{S(a_1S^5 + a_2S^4 + a_3S^3 + a_4S^2 + a_5S + a_6)} \qquad (172)$$

where a_1, a_2 ... a_7 are defined by equations (146-1) to (146-7). The optimum values of K and h are used in the determination of these coefficients.

The frequency response of equation (172) as simulated using programme FREQRES appears in Fig. 47.

The phase margin is 62°; while the gain margin is 8 db. This implies that the loop gain may be multiplied by a factor of 2.55 before the stability becomes limiting. (In the event of a failure in the adaptive gain compensator, stability will only become limiting under the extreme condition of 65,000 p.s.i. pressure drop across the orifice.)

6.17 Final Considerations

6.17-1 Hydraulic power supply system

The salient features of the hydraulic power supply system have been described in section 6.7. Some of the practical details however, warrant further discussion. The recommended hydraulic circuit appears in Fig. 48, in which the graphical symbols used are to J.I.C. standards.

The fixed displacement pump draws the fluid through filter FLT_1, which has a 100 micron element. The delivery pressure of the pump is limited by relief valve, V_3. The sequence valve, which is supplied via a 20 micron filter guarantees flow to the servo-valve circuit, which is connected to the primary of the sequence valve. The pressure to the servo-valve is reduced using reducing valve V_5. The servo-valve is protected against faulty reducing valve behaviour, or pressure surges in the lines by relief valve V_4, which should be a single stage type for rapid response. The servo-valve is protected against contamination, and silting by the 5 micron filter, FLT_3.

The secondary of the sequence valve supplies the HPA circuit. For I.P. operation of the control system, the solenoid of the 2-position, 3-way valve, V_2 would be energised, and the intensifier would be bypassed.

For H.P. operation, V_2 is de-energised and the accumulator is charged by the intensifier. When the accumulator piston reaches the end of its stroke, it opens the limit switch, which de-energises the solenoid of unloading valve V_1, thus unloading the secondary of the sequence valve and decreasing the pump effort. The unloading valve actuator is 'fail-safe' in operation.

The filters should have beck pressure sensitive indicators which provide a measure of the dirt accumulated. The state of the filters should be inspected regularly.

The hydraulic circuit plumbing should be carried out according to J.I.C. standards, and normal maintenance procedures recommended by J.I.C. should be enforced.

6.17-2 Precautions and safety measures

The lengths of the hydraulic lines connecting the servo-valve to the HPA, and the latter to the chamber should be kept as short as possible in order to minimise additional delays in the response.

In estimating the chamber gain and the referred bulk modulus, the air content of the chamber was neglected (sections 6.5 and 6.6). At a chamber pressure of 50,000 p.s.i., by equations (8) and (10.1), the referred bulk modulus would be lowered by about 2% if the fluid volume is 1% air. At lower pressures the reduction in referred bulk modulus would be much greater. It is recommended that the air be evacuated from the chamber immediately before commencement of the pressing cycle. This may be done using an exhaust pump which would form part of the auxiliary equipment and controls the press.

It is recommended that the lid of the H.P. chamber be secured to the chamber flange using 6 radial clamps which are equally spaced around the periphery of the flange. The clamps should be so designed that any three, spaced 120° apart could support the full load. The clamps will be self-locking if the flanges and the working faces of the clamps are tapered to an included angle of slightly less than twice the angle of friction. Each clamp may be driven by a pneumatic jack (see Fig. 49) which need only apply a small force (of the order 50 to 100 lb_f) during clamping. The arrangement may be protected by 'fail-safe' limit switching which ensures that the pressurisation of the chamber cannot occur before the lid is clamped.

The H.P. chamber should be protected against overpressures due to control system failure, by a safety valve (see section 3.2). Furthermore, the strains on the outer wall of the chamber should be monitored by a strain gauge bridge circuit, which could signal the electronic programmer to discharge the pressure should the strains reach undesirably high levels.

6.17-3 Improvements to the mathematical model of the control system

The time constants of some of the elements in the control system (e.g. the pressure transducer, displacement transducer and adaptive gain compensator elements) have all been neglected as they are very small. For cases where these time constants are not negligible, or when greater accuracy is required, the lag terms due to these elements should be incorporated in the overall system transfer function. With the digital computer facilities currently available, there is virtually no limit to the order of transfer function that may be processed using the methods discussed in the preceding sections. The increased accuracy should however be weighed against the cost of processing time, which increases rapidly with the order of the transfer function.

Delays caused by flow paths in the control circuit should also be taken into account if greater precision is required.

6.17-4 Improvement of system performance

The accuracy of the output is governed by the accuracy of the pressure transducer in the major loop. A better quality pressure transducer will be required if greater accuracy is necessary.

If the control system is to be used for I.P. operation, some means of compensating for large variations in the chamber gain should be considered. Two methods which could be considered are:-

(i) The incorporation of additional amplifiers in the forward path of the control system, which are switched out of the circuit by signals from the programmer as the I.P. cycle progresses and the cylinder gain rises due to compaction of the powder.

(ii) An adaptive control system which monitors the flow rate into or out of the chamber and the rate of chamber pressure change. This information is used to evaluate the chamber gain, and send a signal to a multiplier in the forward path which offsets any variation.

Further consideration of chamber gain compensation is beyond the scope of this study.

The control system as optimised in the preceding sections does not constitute an "optimal control system" as the term is interpreted in modern control theory [64].

Optimal control theory dictates that all the state variables be fed back so as to minimise a performance index [64]. In the current design, many of the state variables are inaccessible, and although methods of 'modelling' portions of the system (so as to obtain approximations of the inaccessible state variables) have been proposed [64], it is believed that the improvement in performance that may be gained by adopting such procedures, will not justify further complication of the control system.

7. DISCUSSION

7.1 General

Continuous pressure control systems of type described in Chapters 4 and 5 have been used with every success in sophisticated presses employed in commercial diamond synthesis [66]. These control systems are simple and reliable, but somewhat less efficient than the type described in Chapter 6.

The basic approach used for continuous pressure control in the H.P. range, i.e. the incorporation of an additional hydraulic amplifier stage in the system may be applied equally well to control applications in the L.P. and I.P. ranges. The benefits of improved efficiency should however be weighed against the disadvantages of complicating the control systems.

The methods of analysis and optimisation discussed in Chapter 6 may be applied to the L.P. and I.P. control systems as well.

The pressure range of the basic control system developed for H.P. use in Chapter 6 may easily be extended to the V.H.P. region. Intensifiers capable of delivering up to 1 c.i.s. of fluid at pressures of up to 150,000 p.s.i. are currently commercially available. For higher pressures, an additional booster stage could be designed for operation in conjunction with standard equipment. Alternatively, a suitable V.H.P. intensifier could be designed and manufactured for pressures in excess of 150,000 p.s.i.

Crossland et al [46] describe a unit suitable for pressures of up to 200,000 p.s.i.

The HPA may also be designed for V.H.P. operation, and is not likely to be the factor which imposes a limit on the pressure level.

The most serious limitation to operating pressures in the upper V.H.P. range is the problem of the tendency of organic fluids to freeze at elevated pressures. The working pressure range of many fluids may be considerably extended by heating the fluid (the tendency for viscosity to increase with increasing pressure is counteracted by its tendency to decrease with increasing temperature). This practice is however restricted to temperatures within the limit of chemical stability of the fluid.

For the upper V.H.P. range, liquid metals show the most promise for utilisation as hydraulic fluids. For example, a range of sodium-potassium eutectoids have been developed for application to nuclear reactor heat-exchangers, and are being considered for utilisation as hydraulic fluids [54].

Sodium-potassium eutectoids have the following properties which make them desirable for utilisation as hydraulic fluids:-

(i) The melting point is usually under 20°F, and the boiling point in excess of 1200°F.

(ii) Very high bulk moduli - of the order 800,000 p.s.i.

(iii) The specific gravity is comparable to that of conventional hydraulic fluids.

(iv) The specific gravity and viscosity are almost invariant with temperature and pressure.

(v) These liquid metals do not freeze until pressures well into the U.H.P. range are reached.

The major difficulties preventing wide scale use of these liquid metals as hydraulic fluids are:-

(i) They oxidise very rapidly.

(ii) They tend to alloy with the materials of hydraulic components, and block orifices and small openings.

The use of sodium-potassium eutectoids is currently limited to preocesses which employ internal pressure intensification. (See sections 2.3 and 6.3-1.)

The computer methods developed for the optimisation of the H.P. system parameters, and for simulating the transient and frequency response are sufficiently general to be applied to a wide variety of control systems. The only limitation is that the control system should be amenable to representation by a transfer function.

For cases where the exact cycle is known, the parameters of the system may be optimised over the complete cycle using an appropriate index of performance. (The ITAE criterion would no longer be applicable, and in the absence of a suitable sophisticated criterion, the ISE or IAE index could be used.)

The methods of continuous pressure control discussed by way of the three examples in Chapters 4, 5 and 6 may be adapted to any of the applications described in Chapter 2. Future developments in the fields of high pressure physics and engineering will undoubtedly give rise to further applications.

7.2 Concluding Remarks

(i) Modern scientific and industrial processes have created a demand for equipment capable of controlling pressure according to predetermined cycles. The demand for such equipment is likely to increase.

(ii) Continuous control of pressure in the L.P. range may be effected with accuracy and reliability using standard components in a simple system the major element of which is the electro-hydraulic servo-valve.

(iii) The operating range of the L.P. System may be extended into the I.P. region using a device such as a proportional pressure reducing valve.

(iv) Continuous control of pressure in the H.P. range may be efficiently effected using a power amplifier stage which is driven by a servo-valve and powered by an intensifier. The performance of the system may be enhanced by providing adaptive gain compensation and by optimising the parameters of the system.

(v) The expression of the transfer function and input to the system in phase variable form facilitates readily programmed digital computer optimisation of the system performance, and simulation of its transient response.

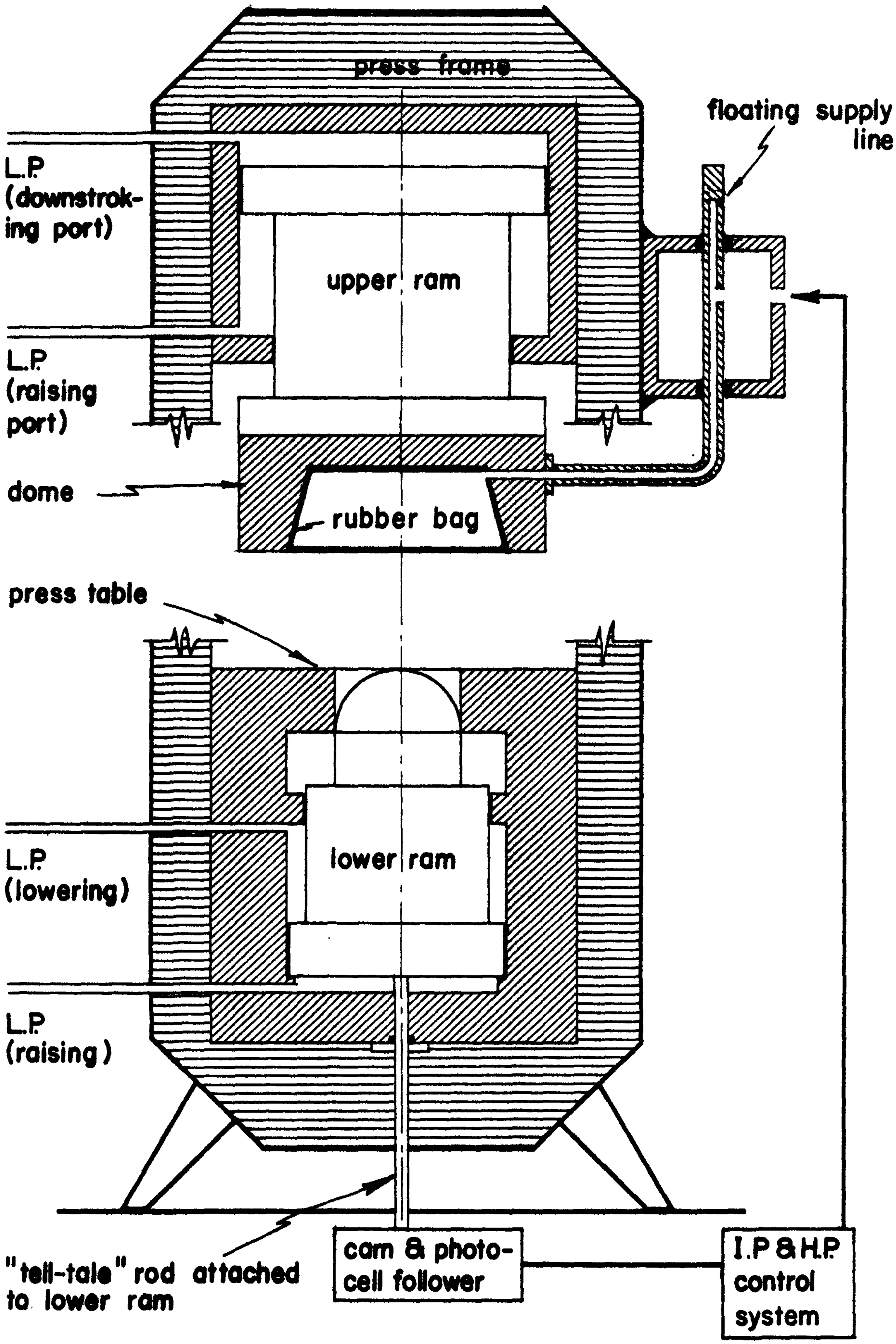


FIGURE 1 Hydrostatic Forming Press

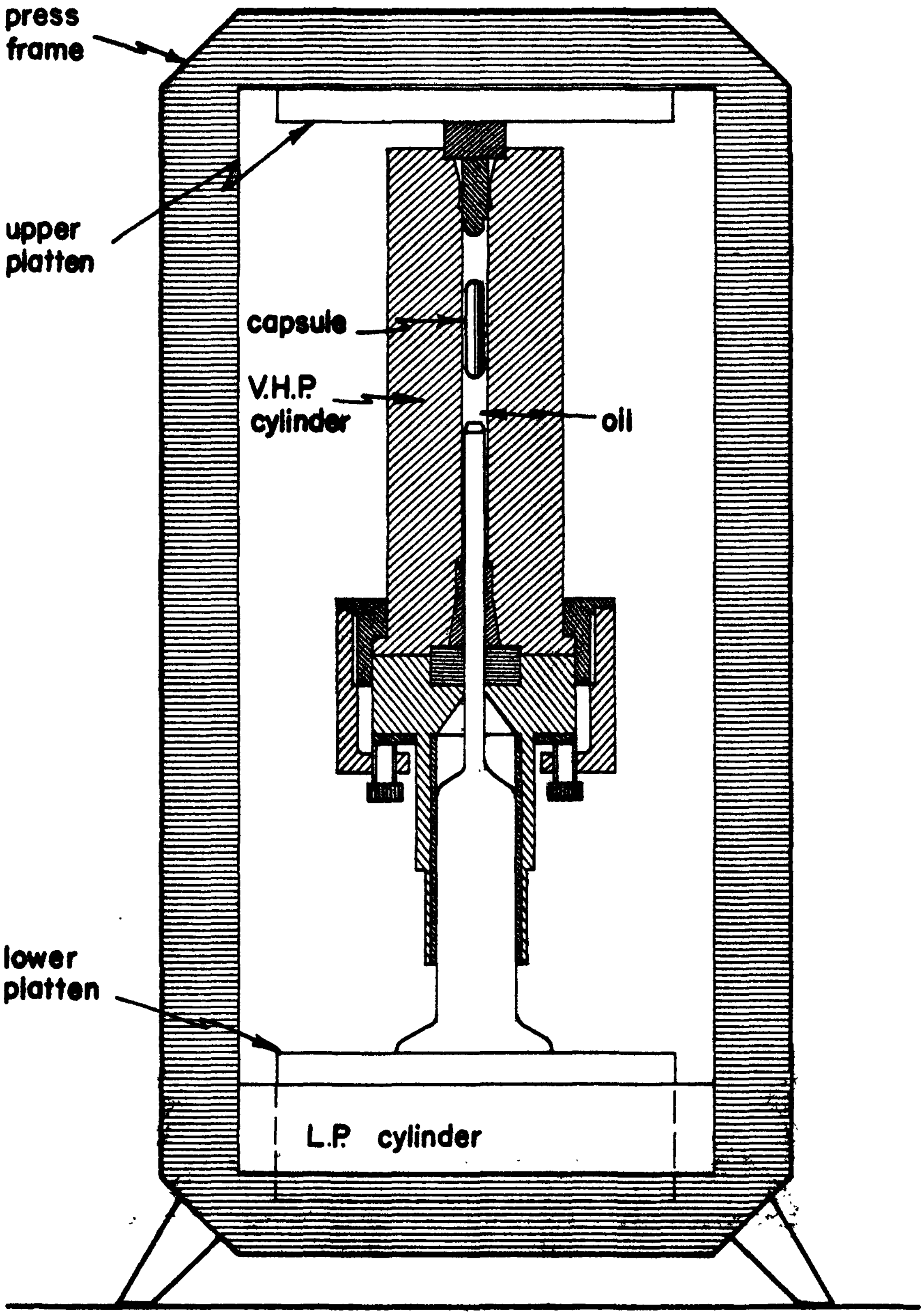


FIGURE 2 Piston-Cylinder Apparatus for H.P & V.H.P Organic Chemical Synthesis.

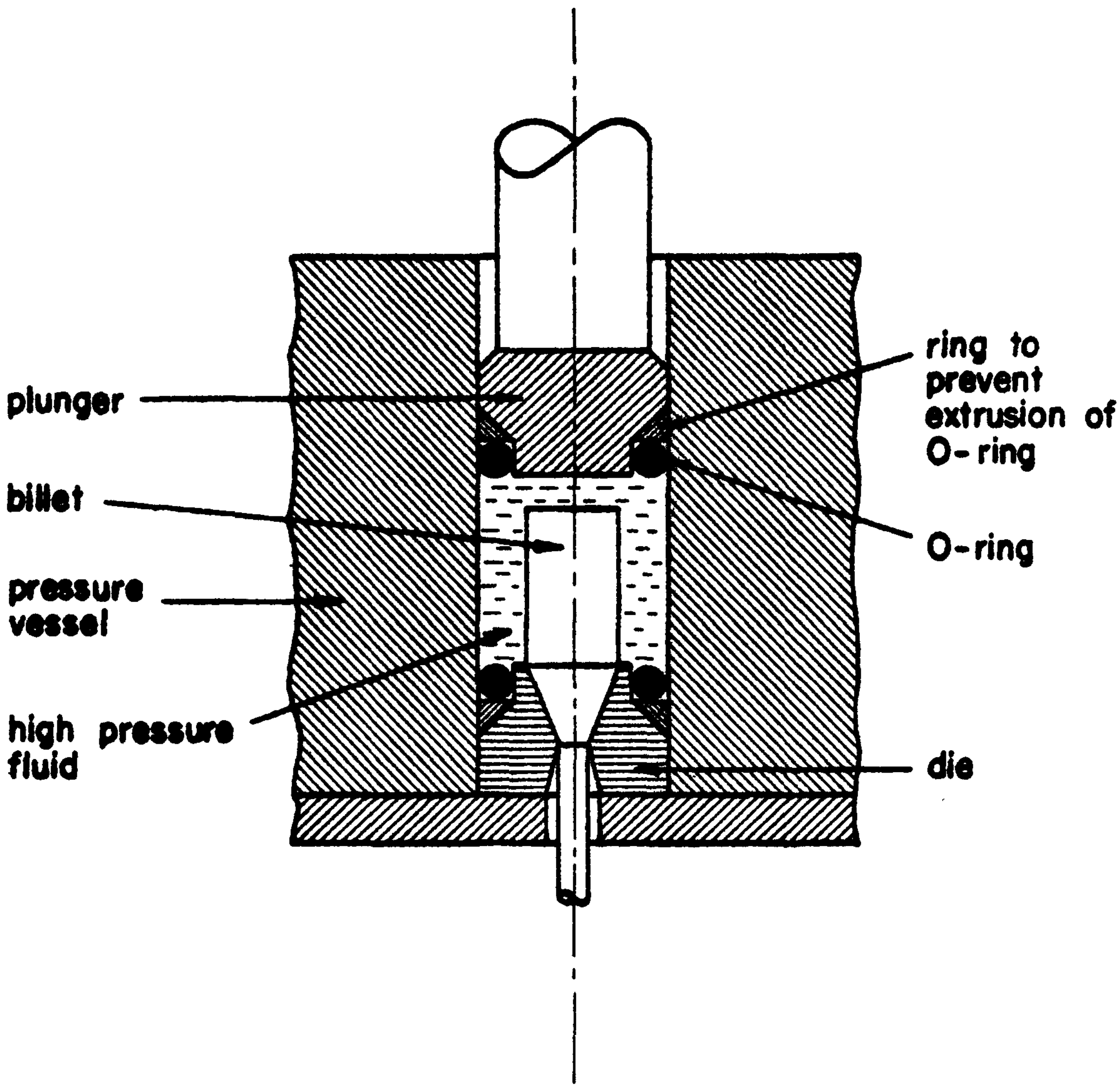


FIGURE 3 Hydrostatic Extrusion

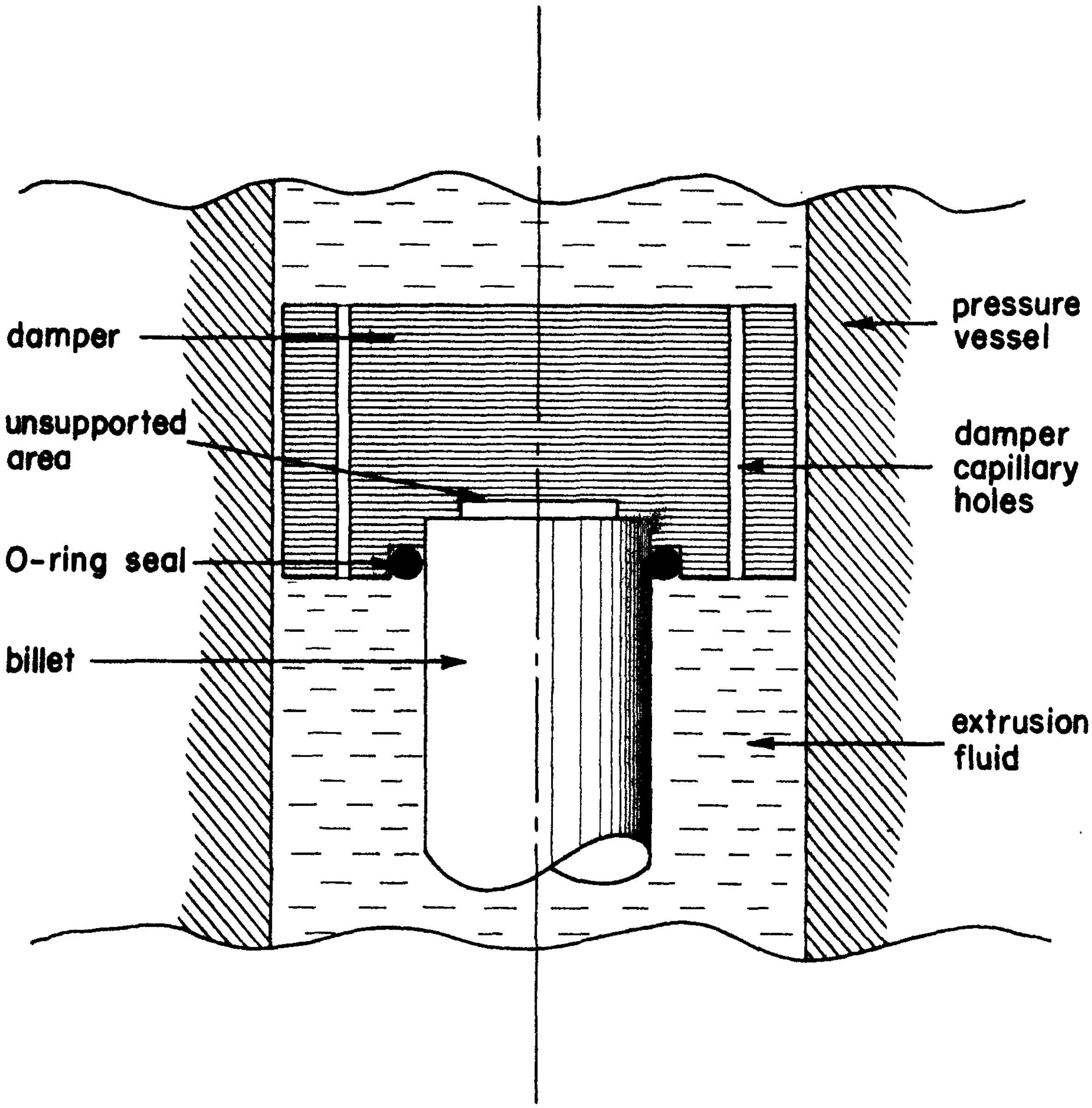


FIGURE 4 Damper Cap for Hydrostatic Extrusion

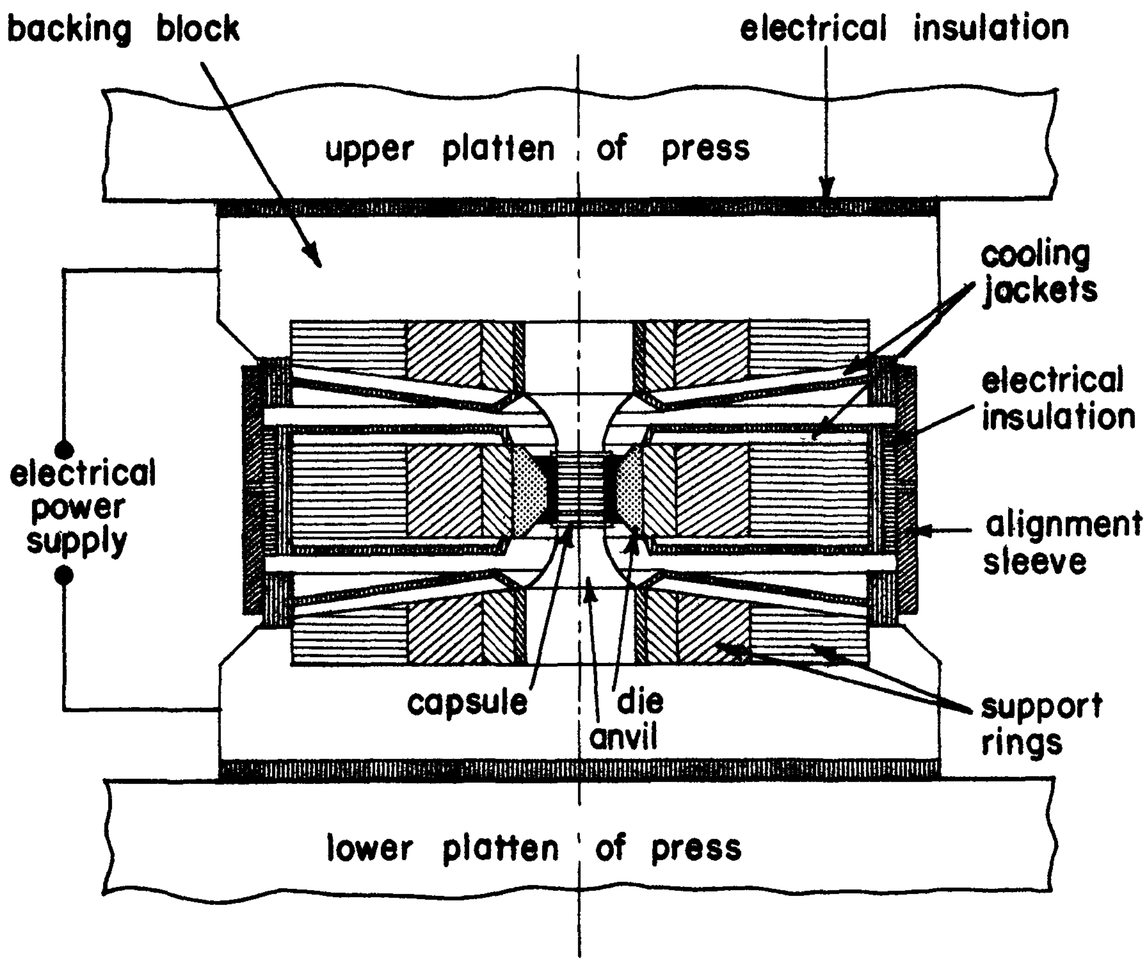
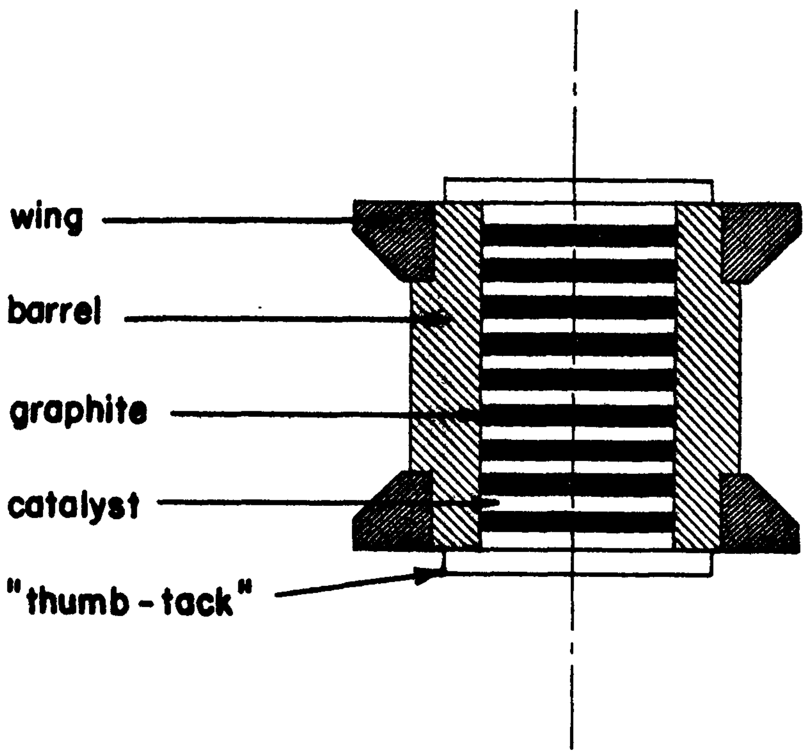


FIGURE 5   A 'Belt-System' Die-Set



<u>FIGURE 6</u> An Assembled Capsule

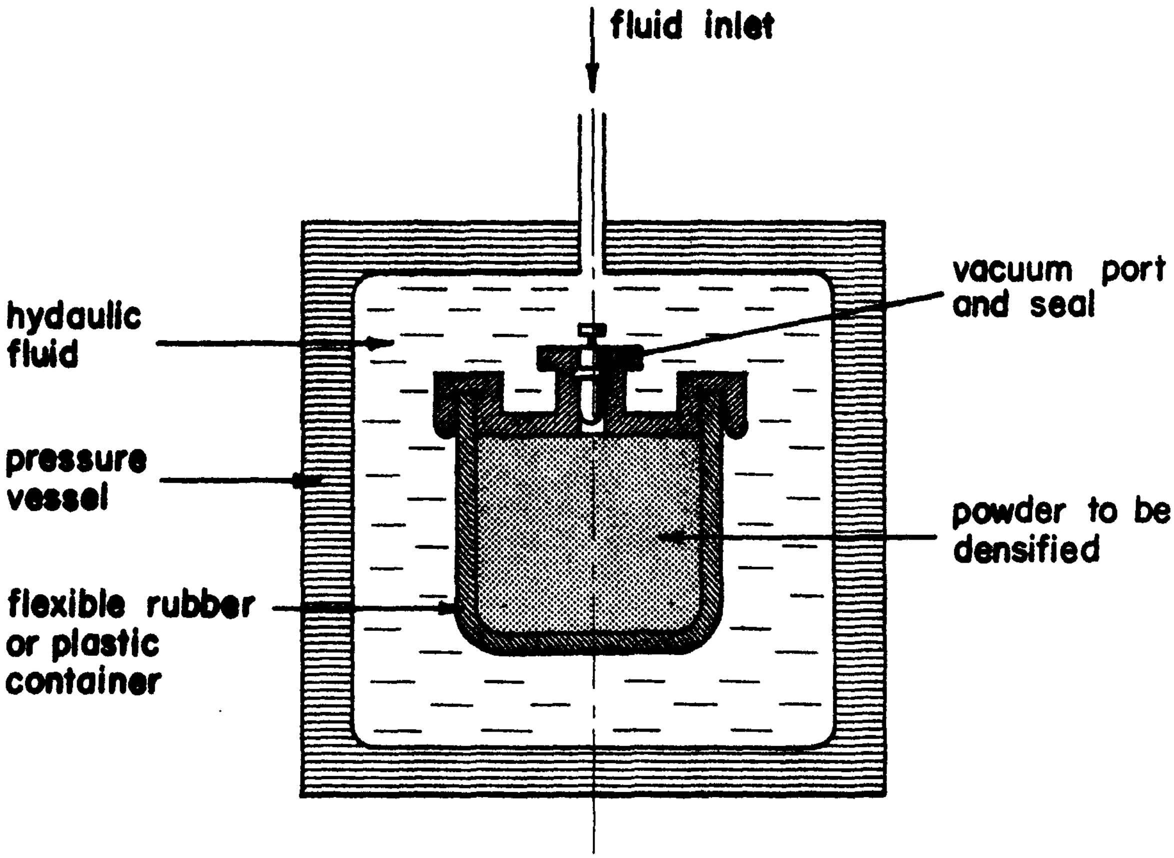


FIGURE 7 Isostatic Pressing

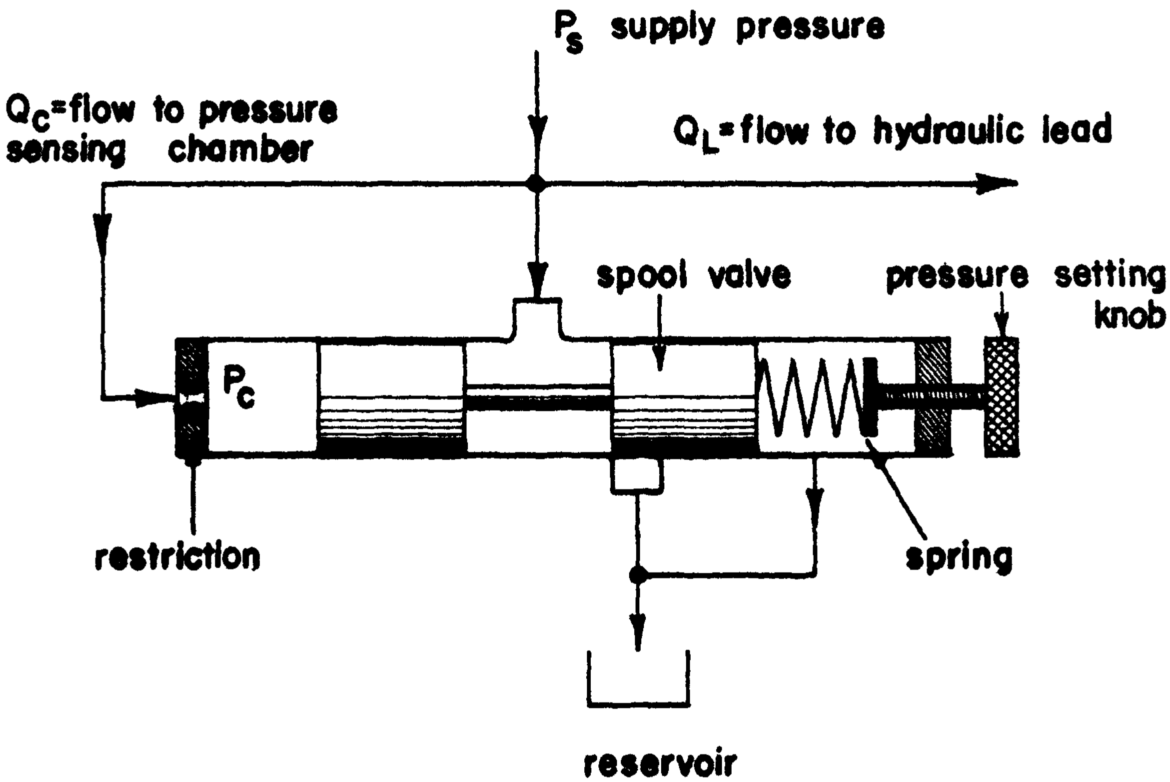


FIGURE 8 Single Stage Relief Valve

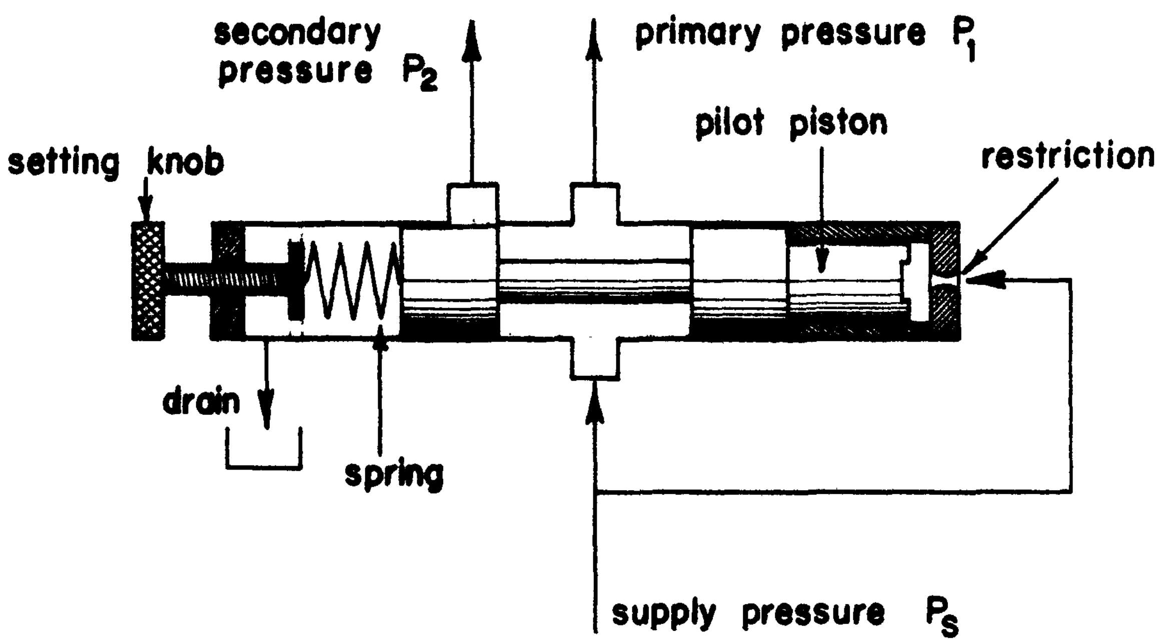


FIGURE 9 A Simple Sequence Valve

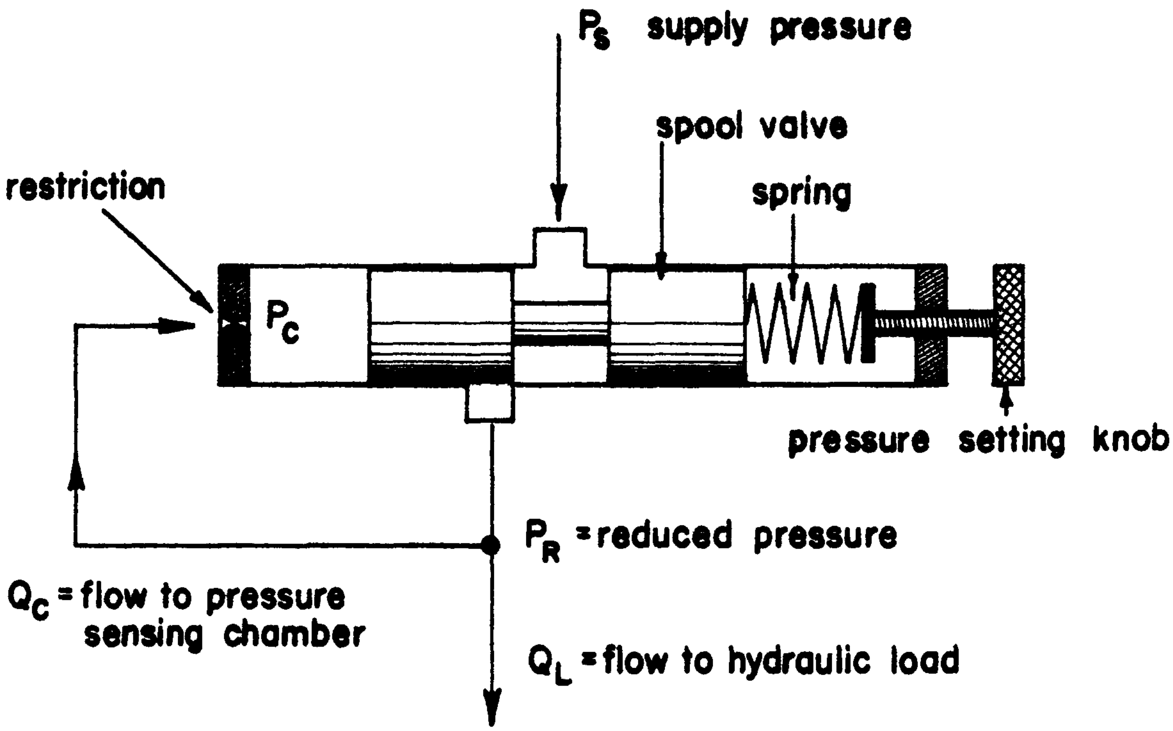


FIGURE 10 Single Stage Reducing Valve

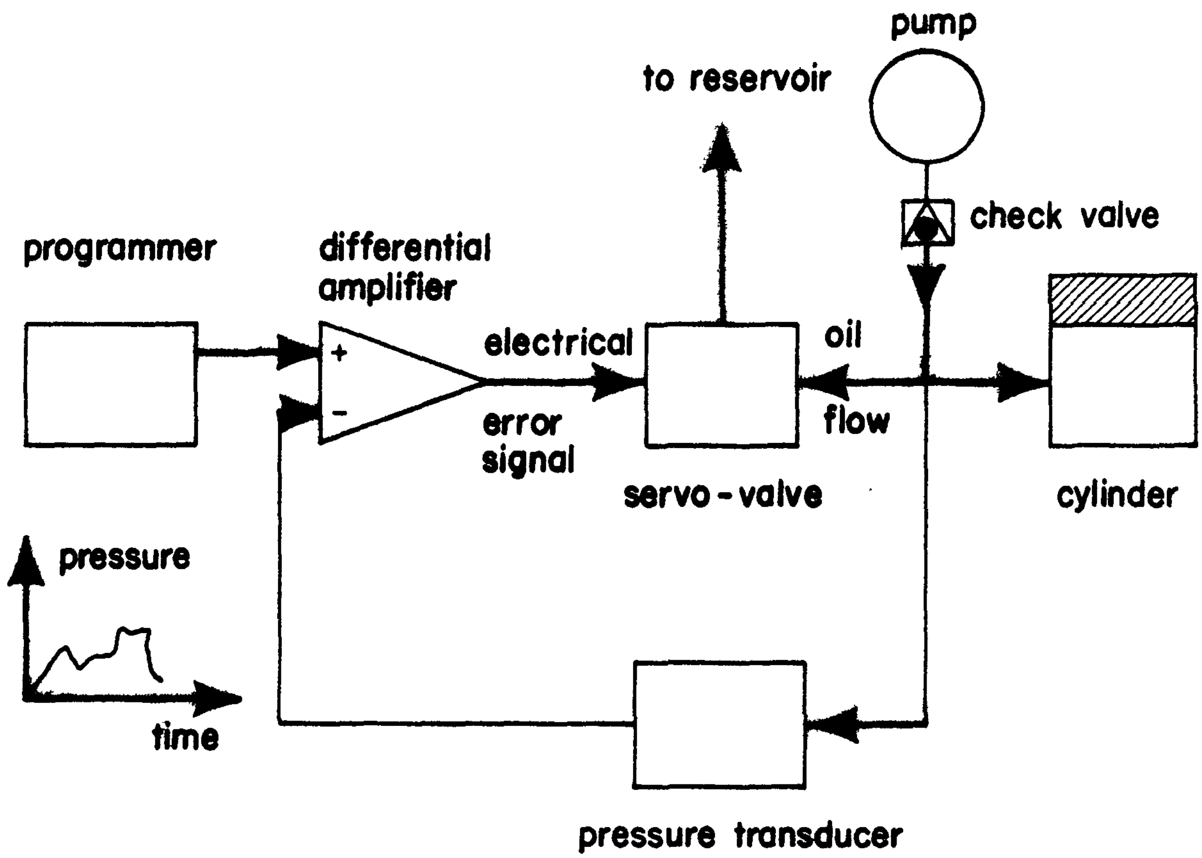


FIGURE 11 Schematic of L.P. Control System

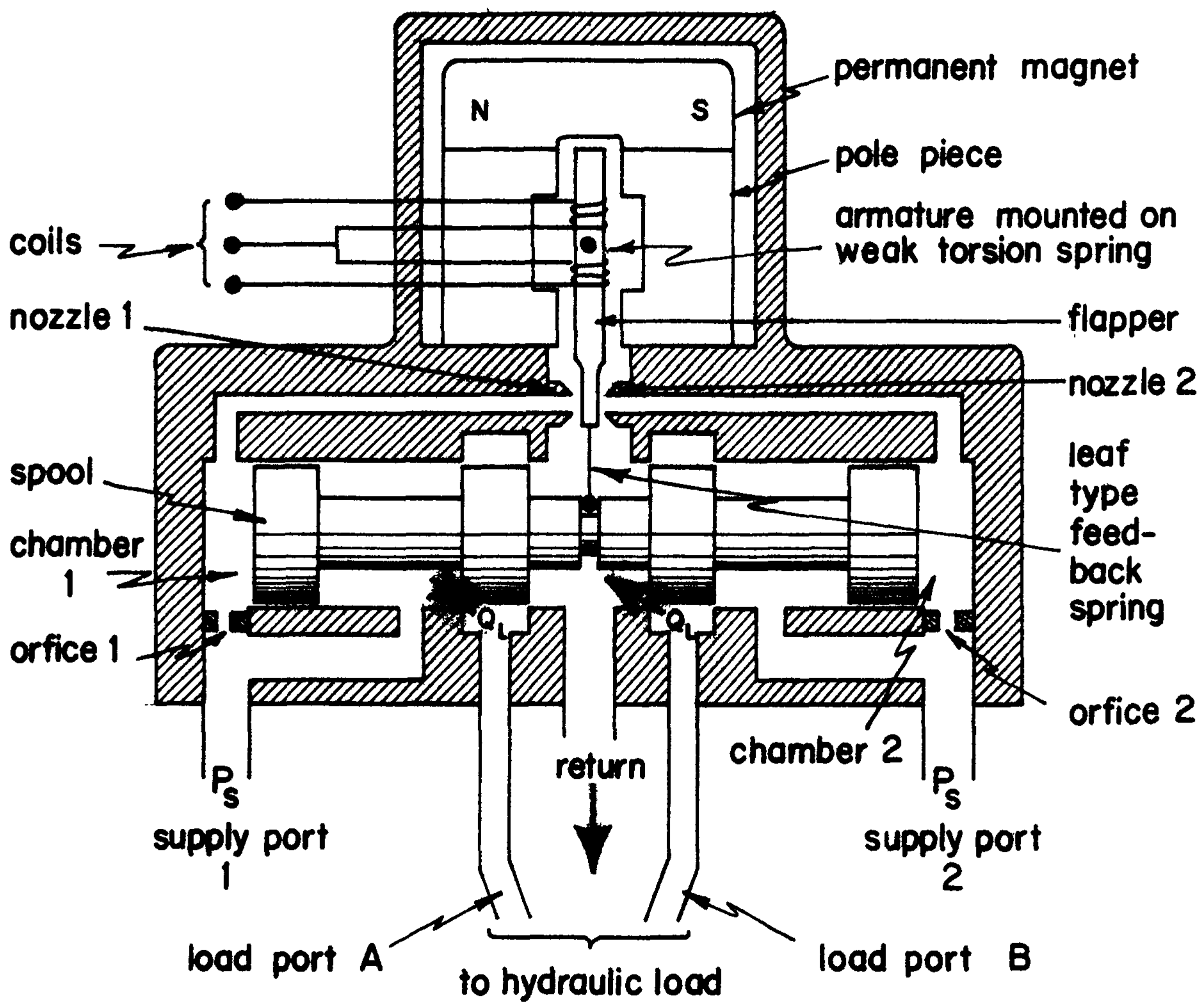


FIGURE 12 Schematic of a Two-stage Electrohydraulic Servo-valve with Position Feed-back

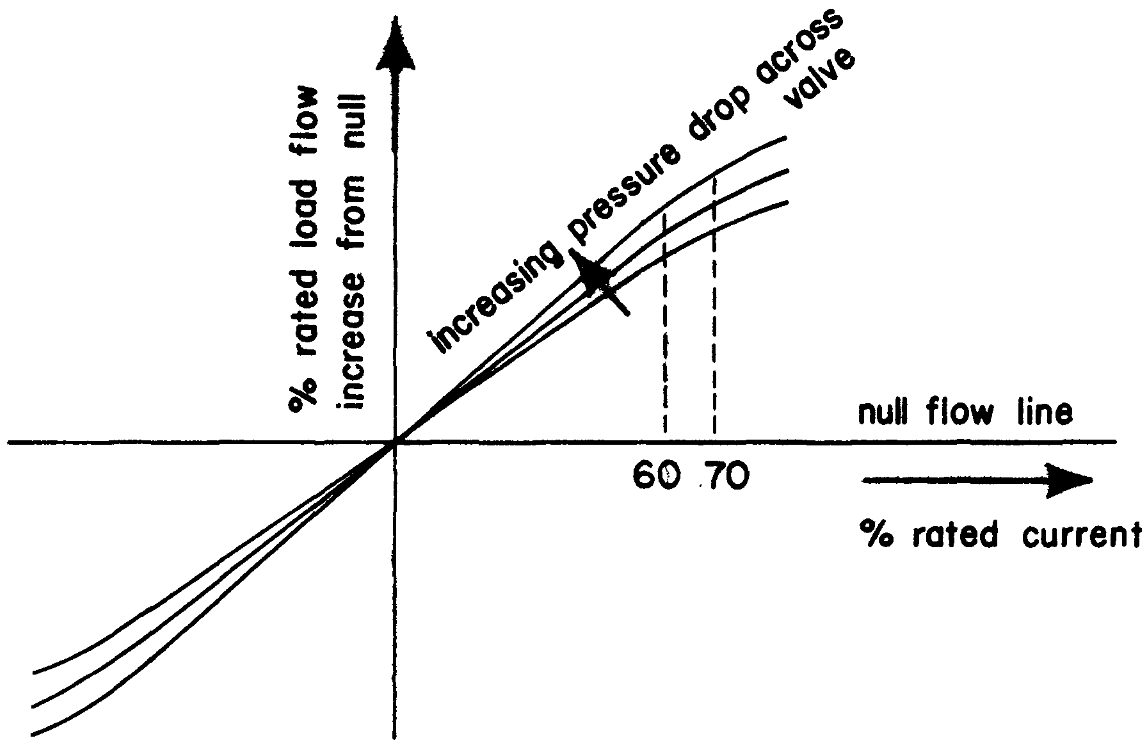


FIGURE 13 Flow-coil Current Characteristics for Typical Servo-valve with Position Feedback

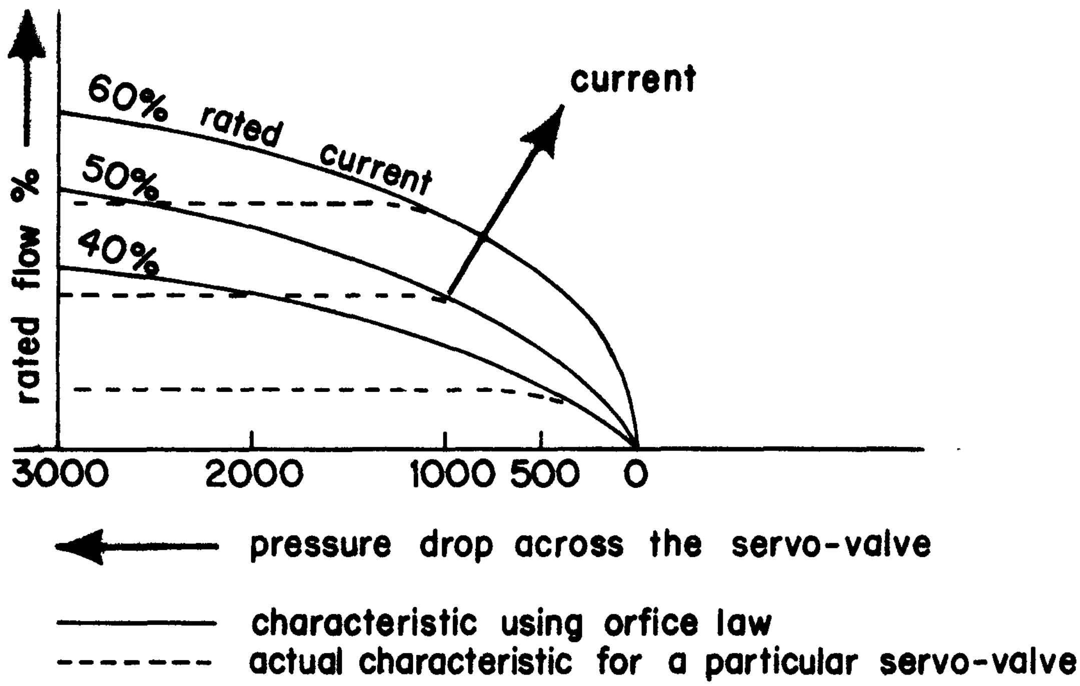


FIGURE 14 Flow-pressure Characteristics for a Servo-valve

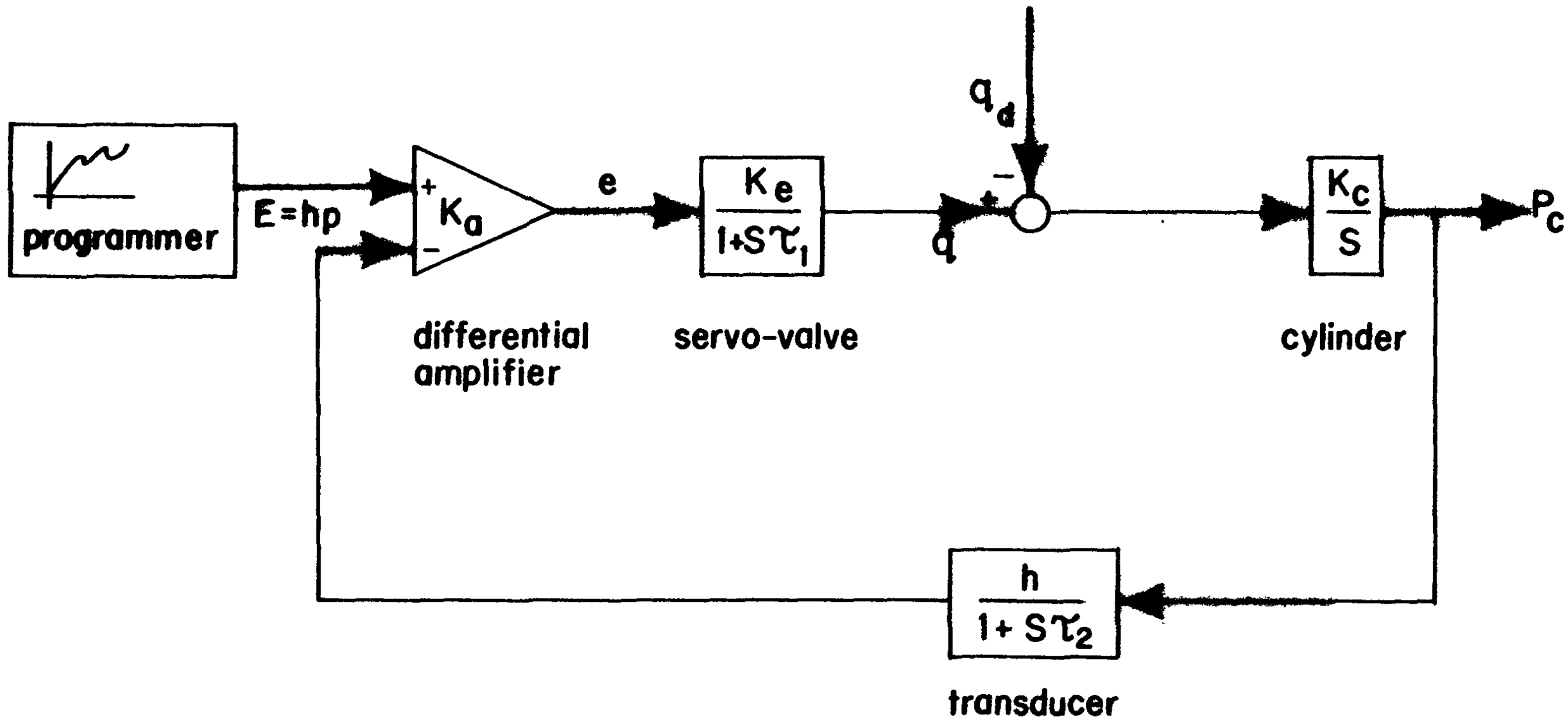


FIGURE 15   Block Diagram of Control System

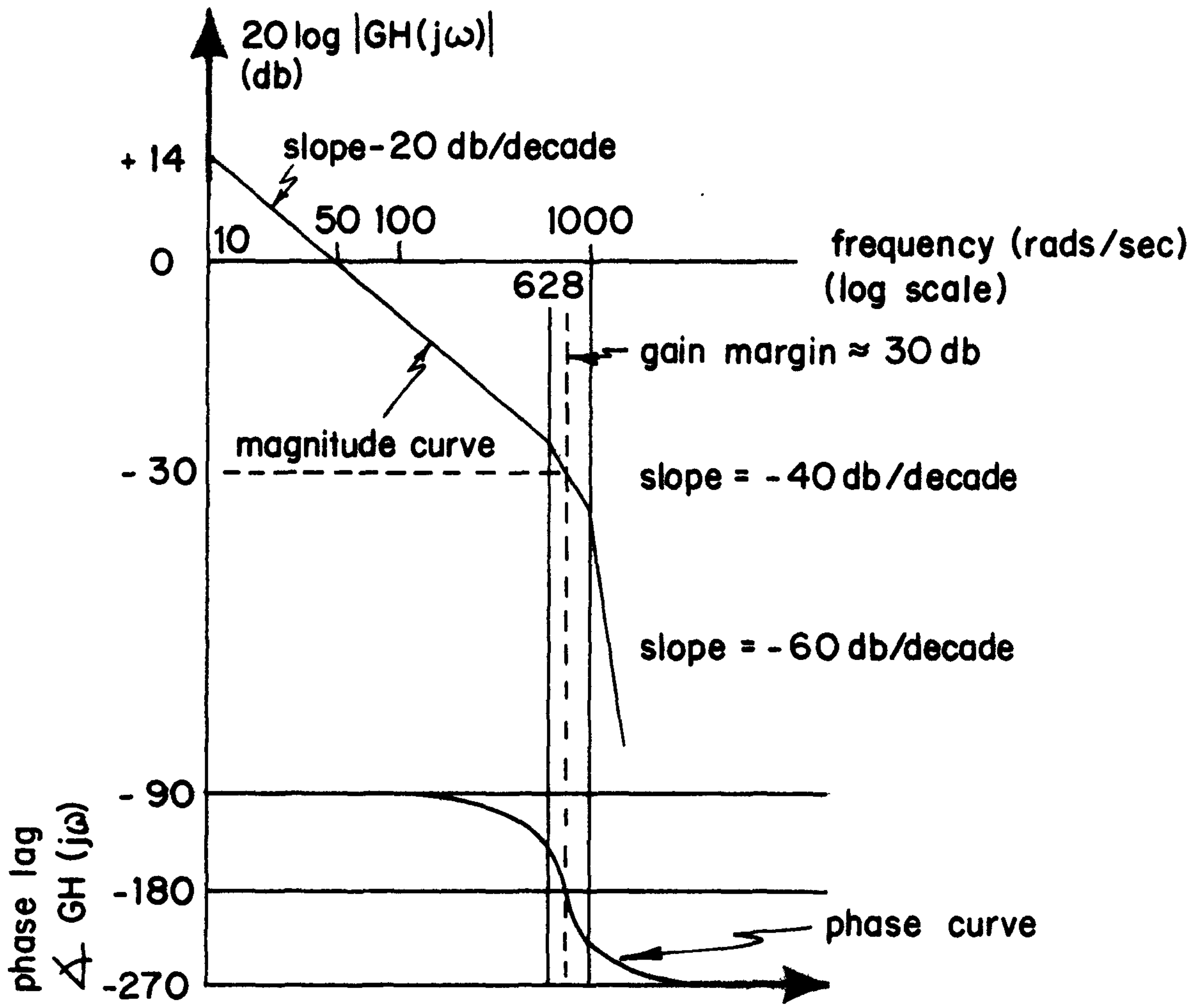


FIGURE 16 Frequency Response of Open-loop System for K = 50

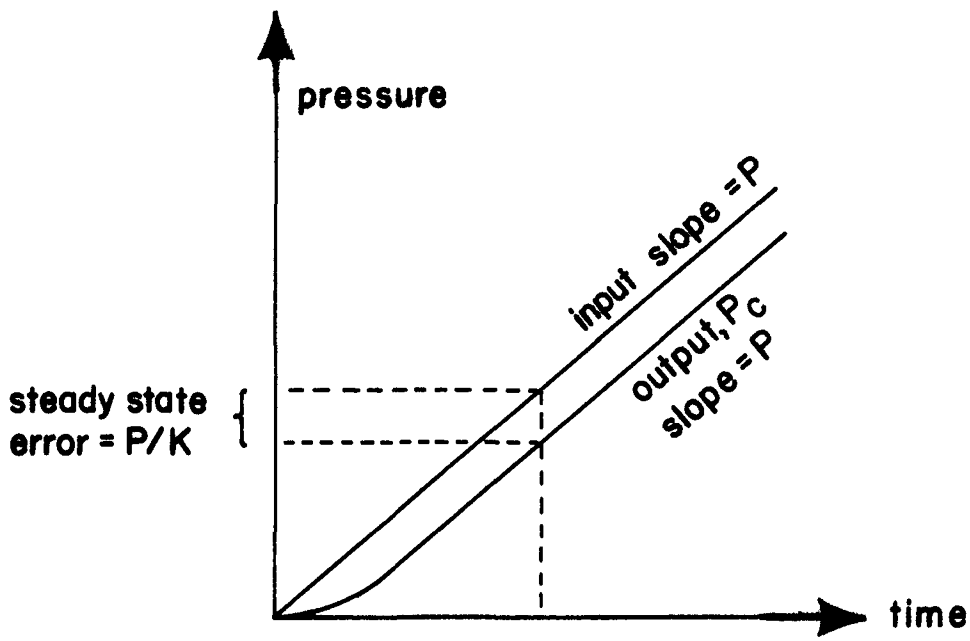


FIGURE 17 Response to Ramp Function

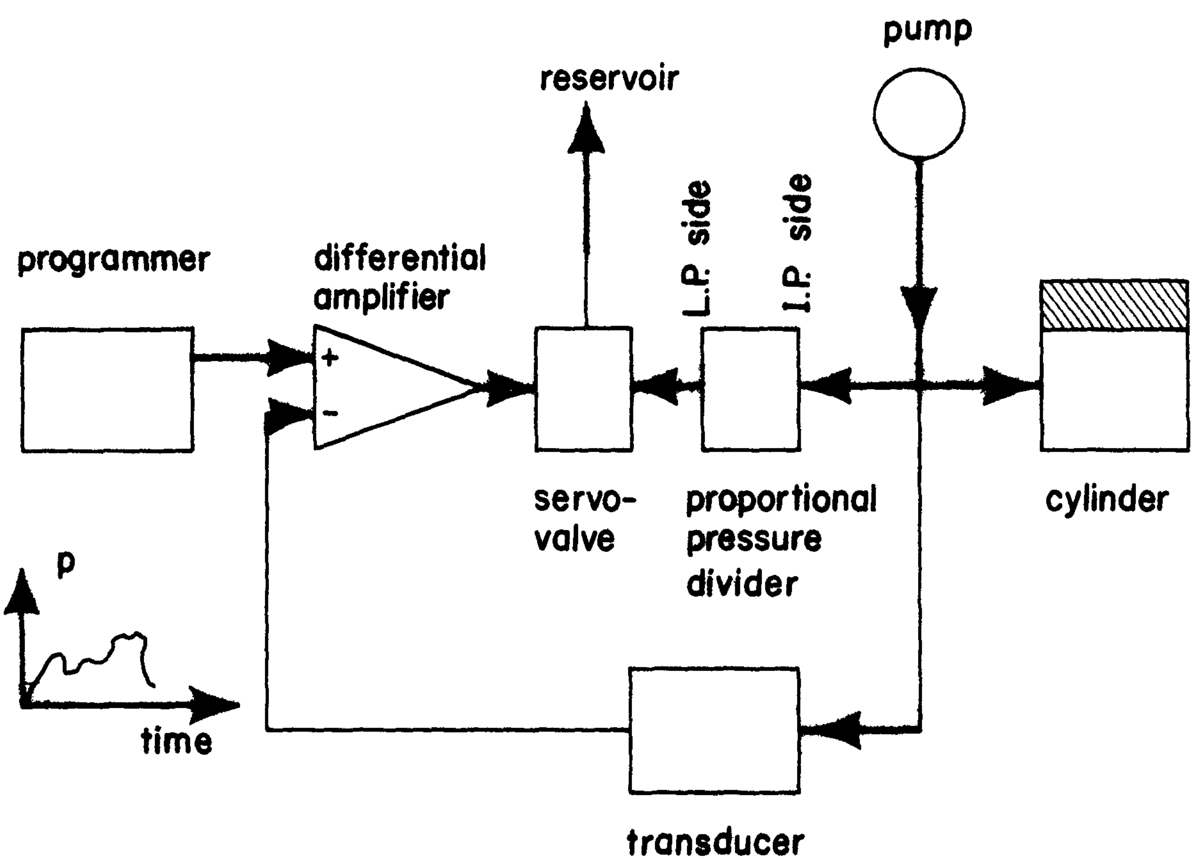


FIGURE 18 Schematic of I.P. Control System

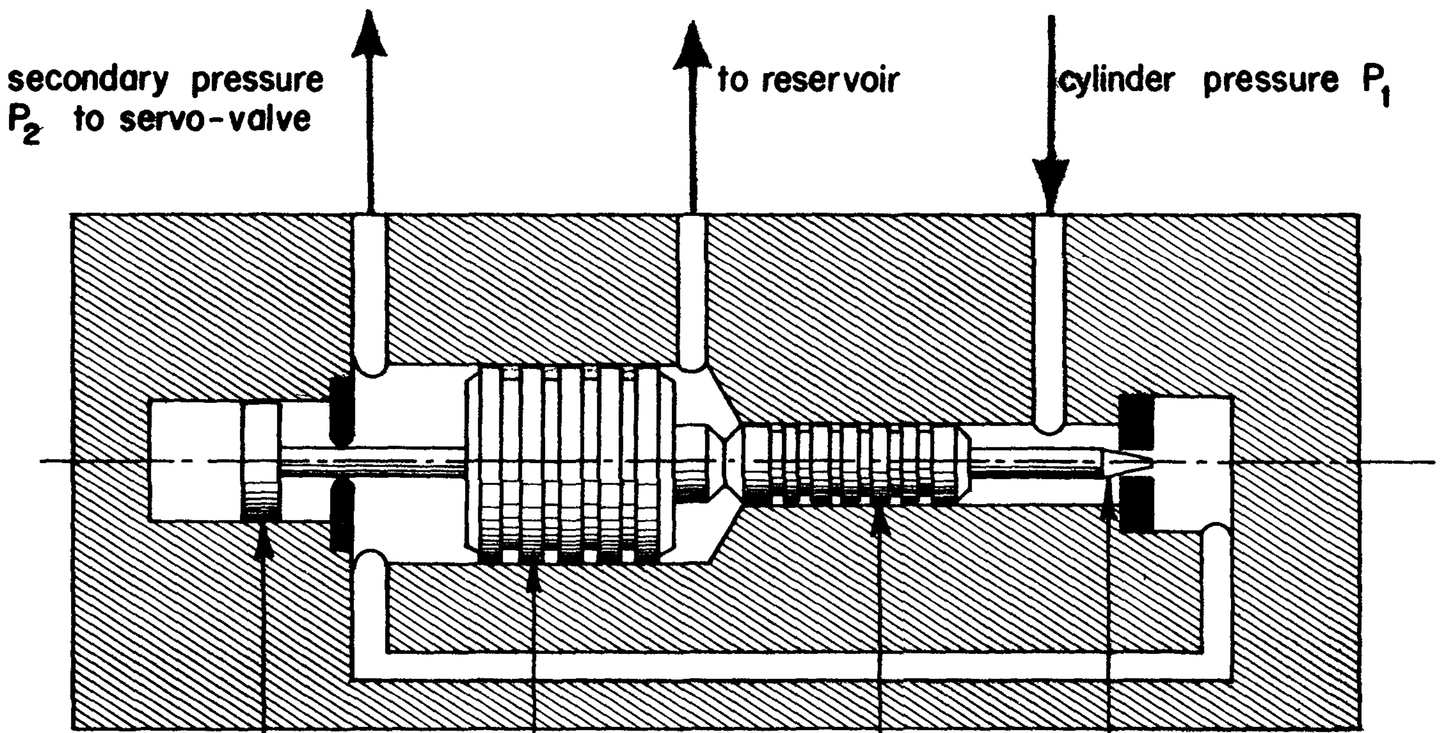


FIGURE 19 Principle of Proportional Pressure Divider

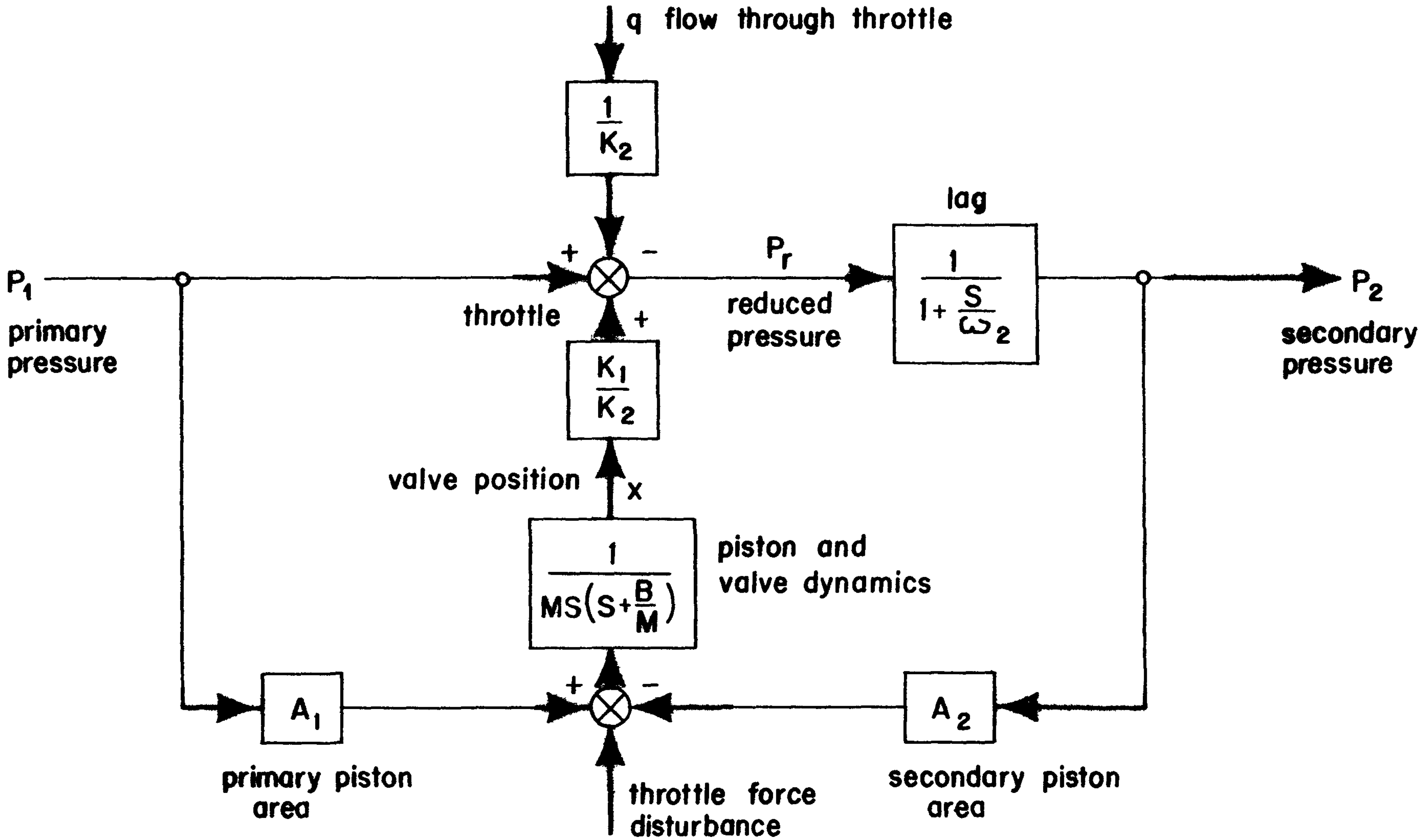


FIGURE 20 Block Diagram of PPD

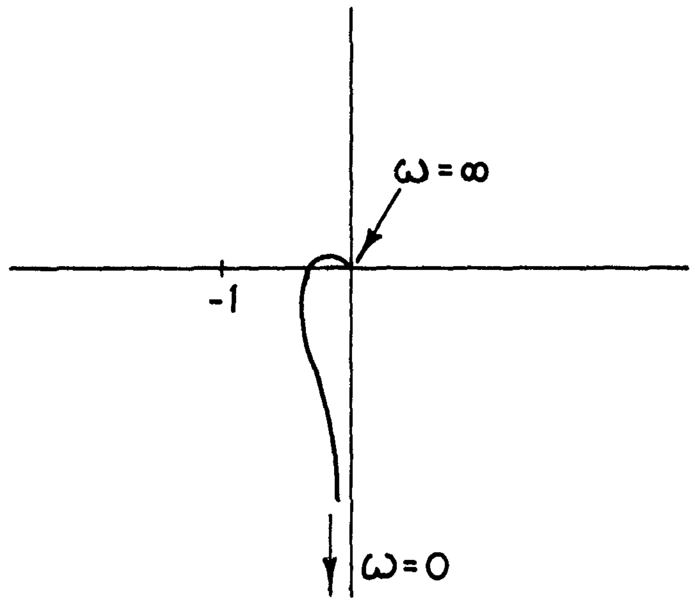


FIGURE 21 (a) Polar Frequency Response Plot of Open-loop Transfer Function with Flow Path Lag

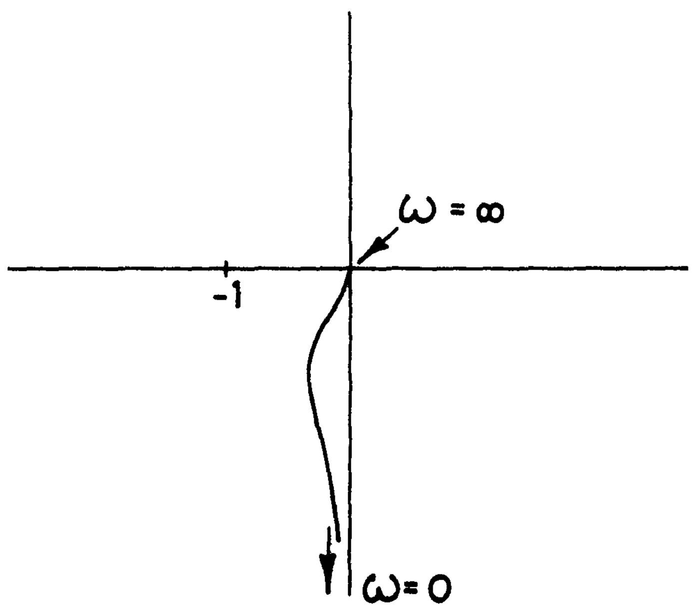


FIGURE 21 (b) Polar Frequency Response Plot of Open-loop Transfer Function After Elimination of Flow Path Lag

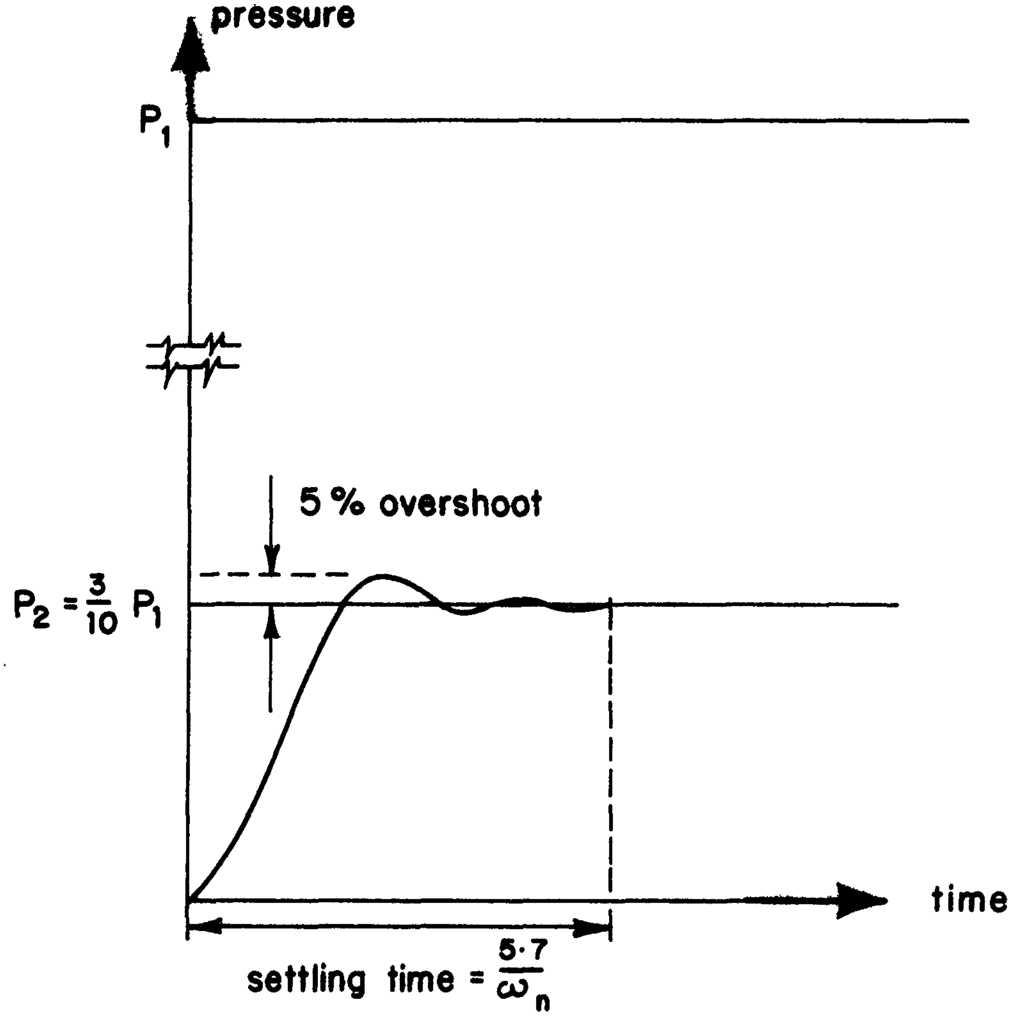


FIGURE 22 Transient Pressure Response of PPD with a Damping Ratio, $\zeta = 0{\cdot}7$

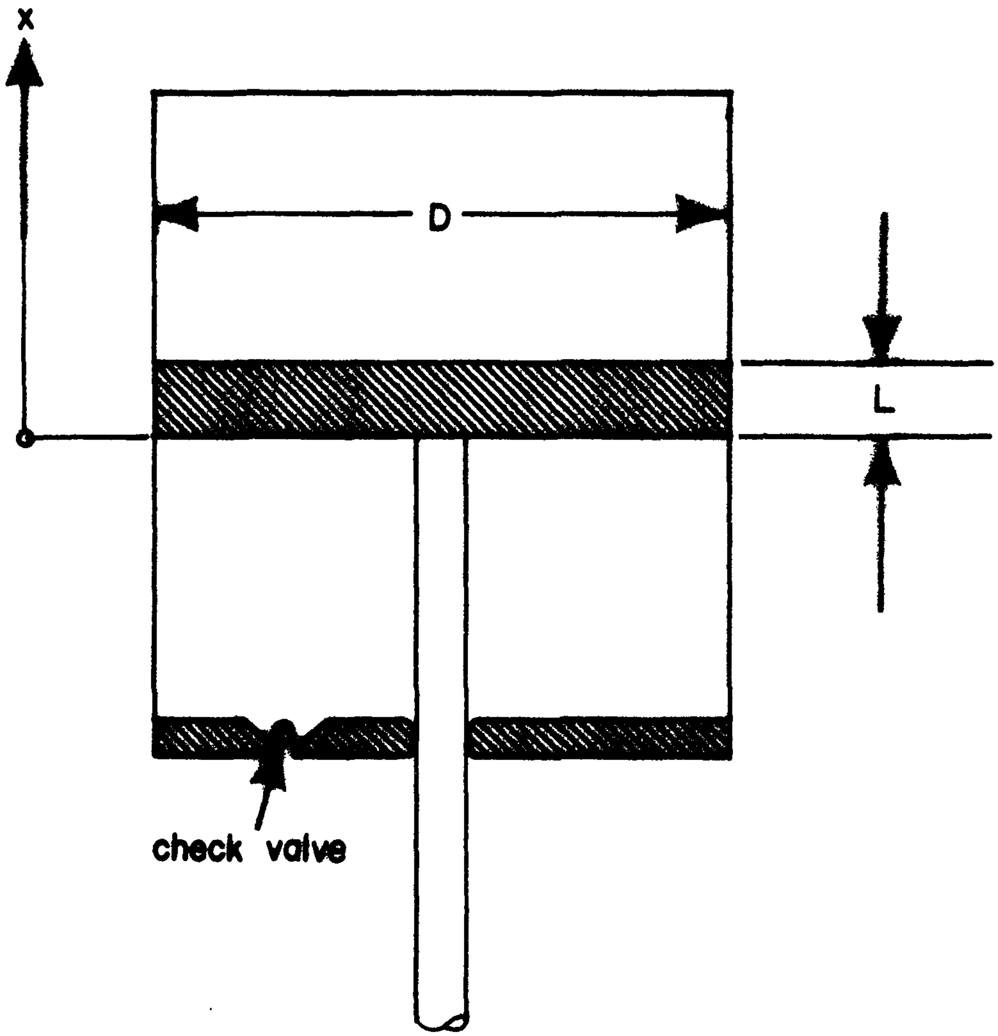


FIGURE 23 Damper

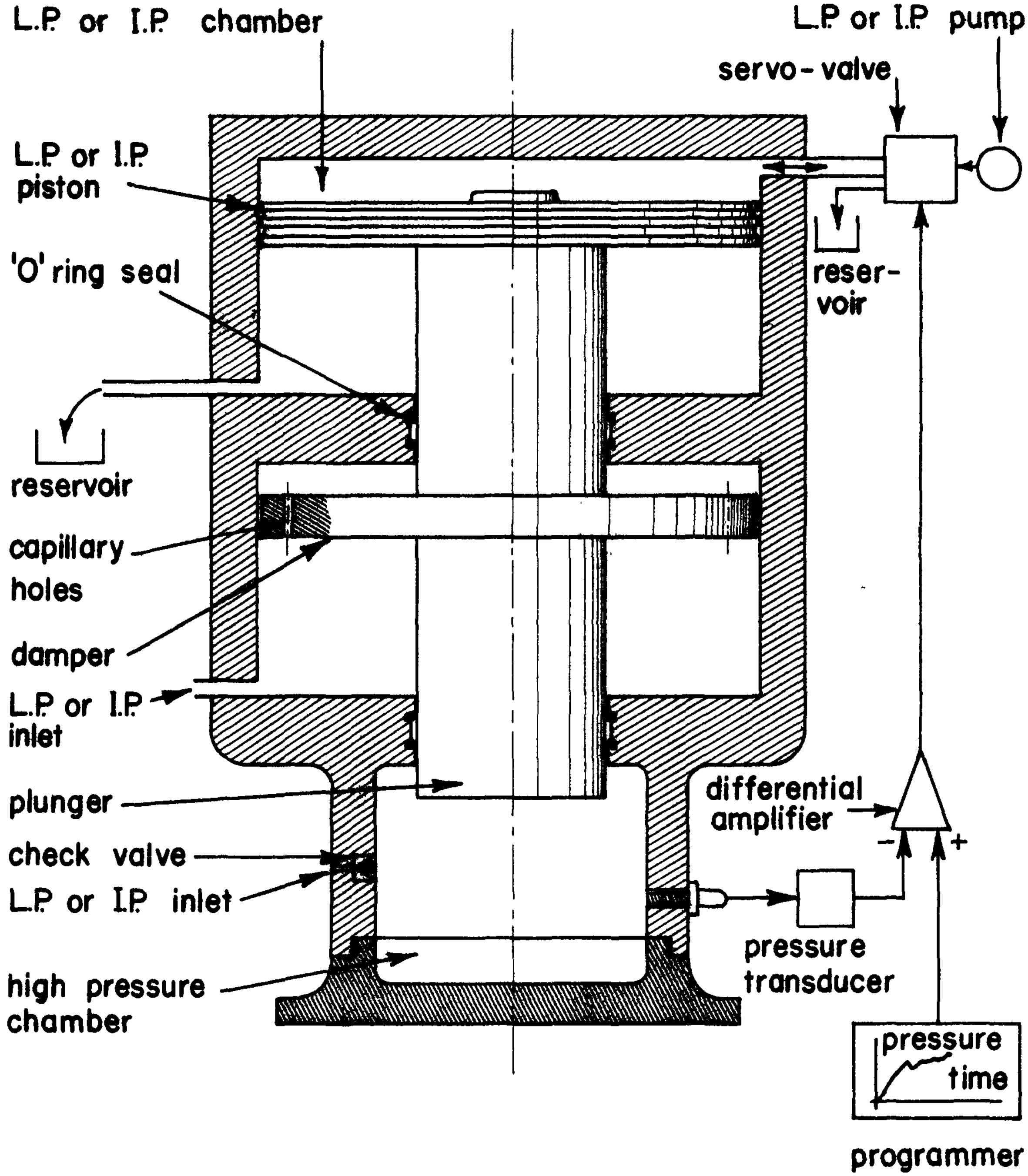


FIGURE 24 Schematic of Pressure Control System for an Isostatic Press with Internal Pressure Intensification

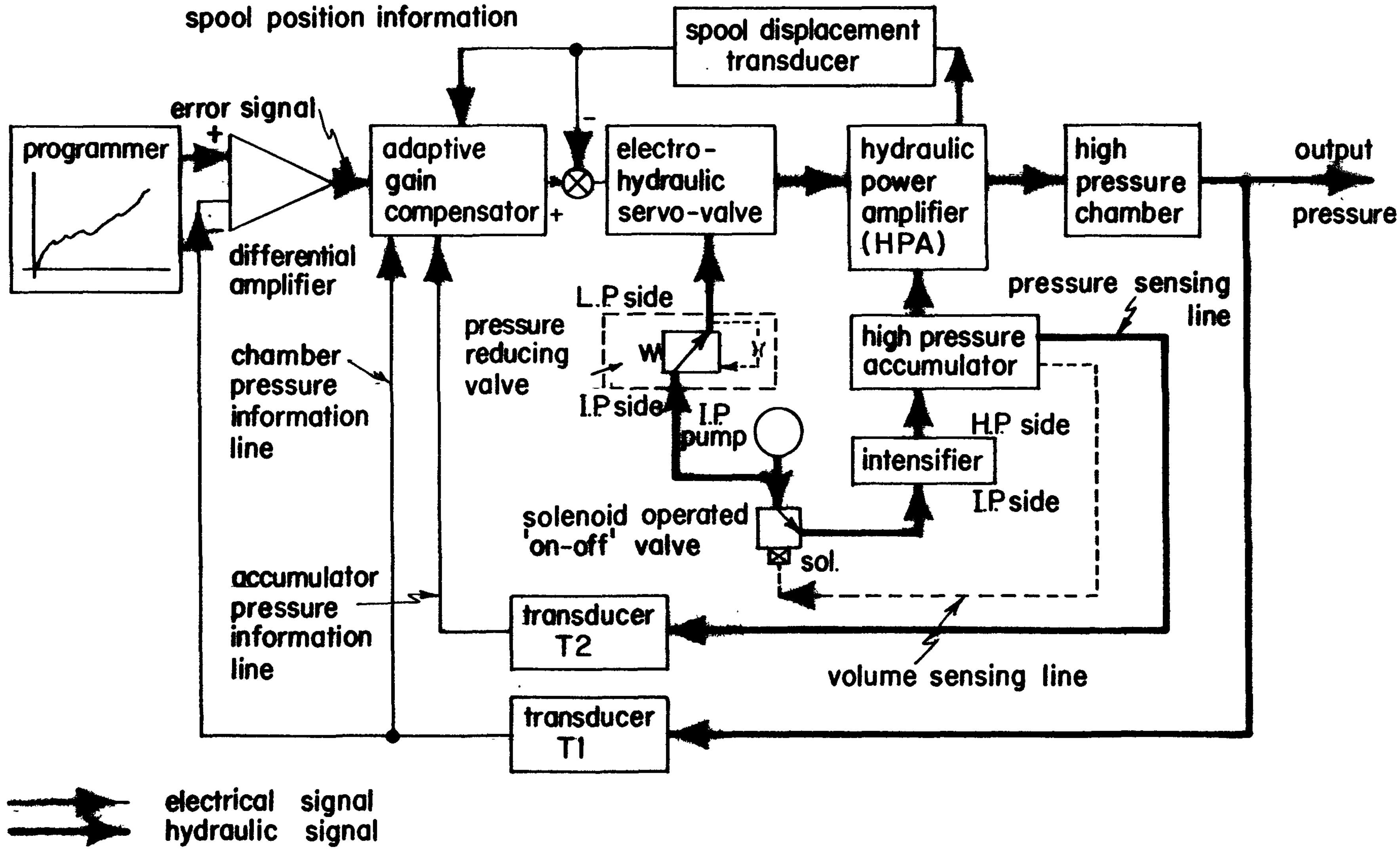


FIGURE 25 Schematic of High Pressure Control System

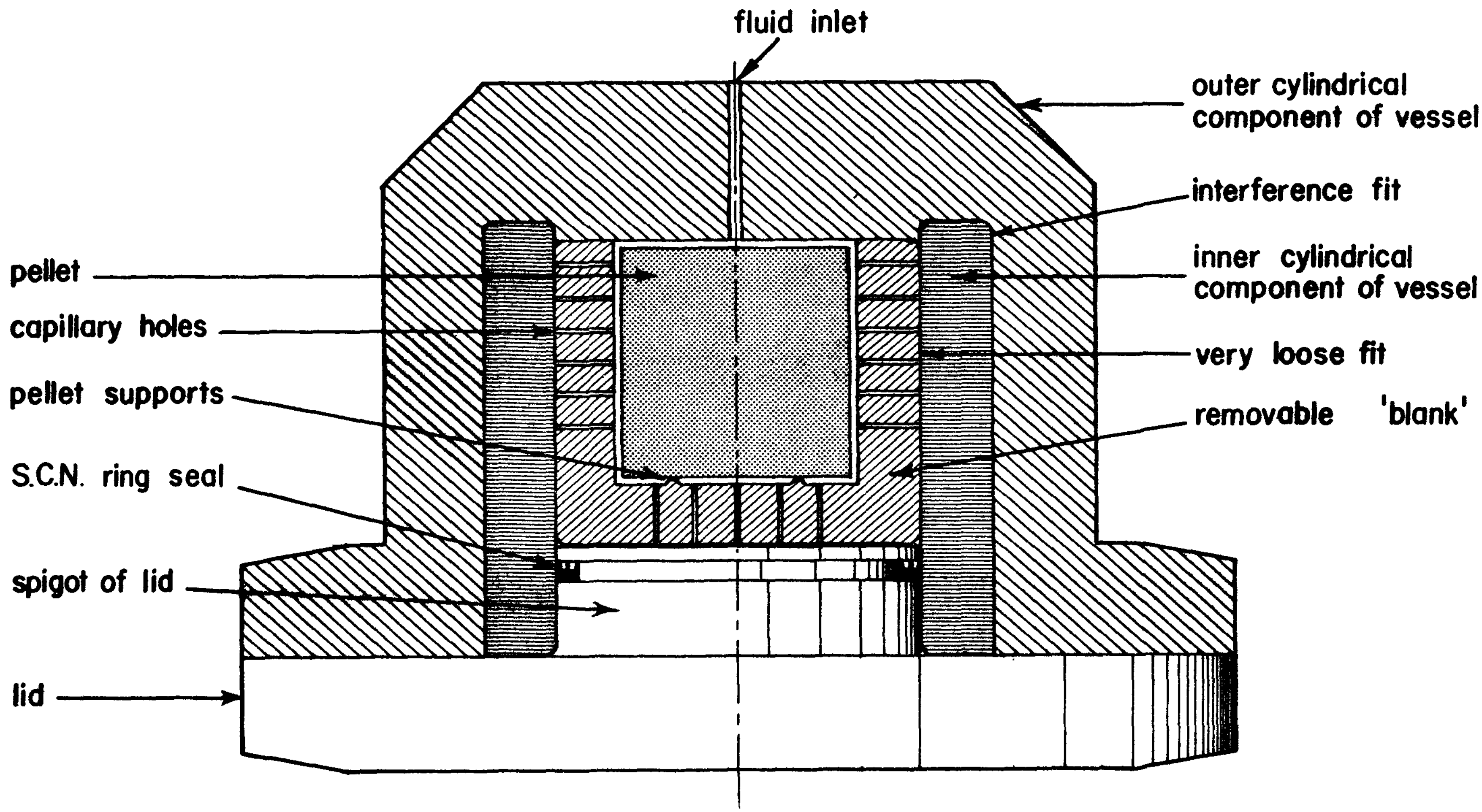


FIGURE 26   Pressure Vessel with Removable 'Blank' for I.P or H.P use.

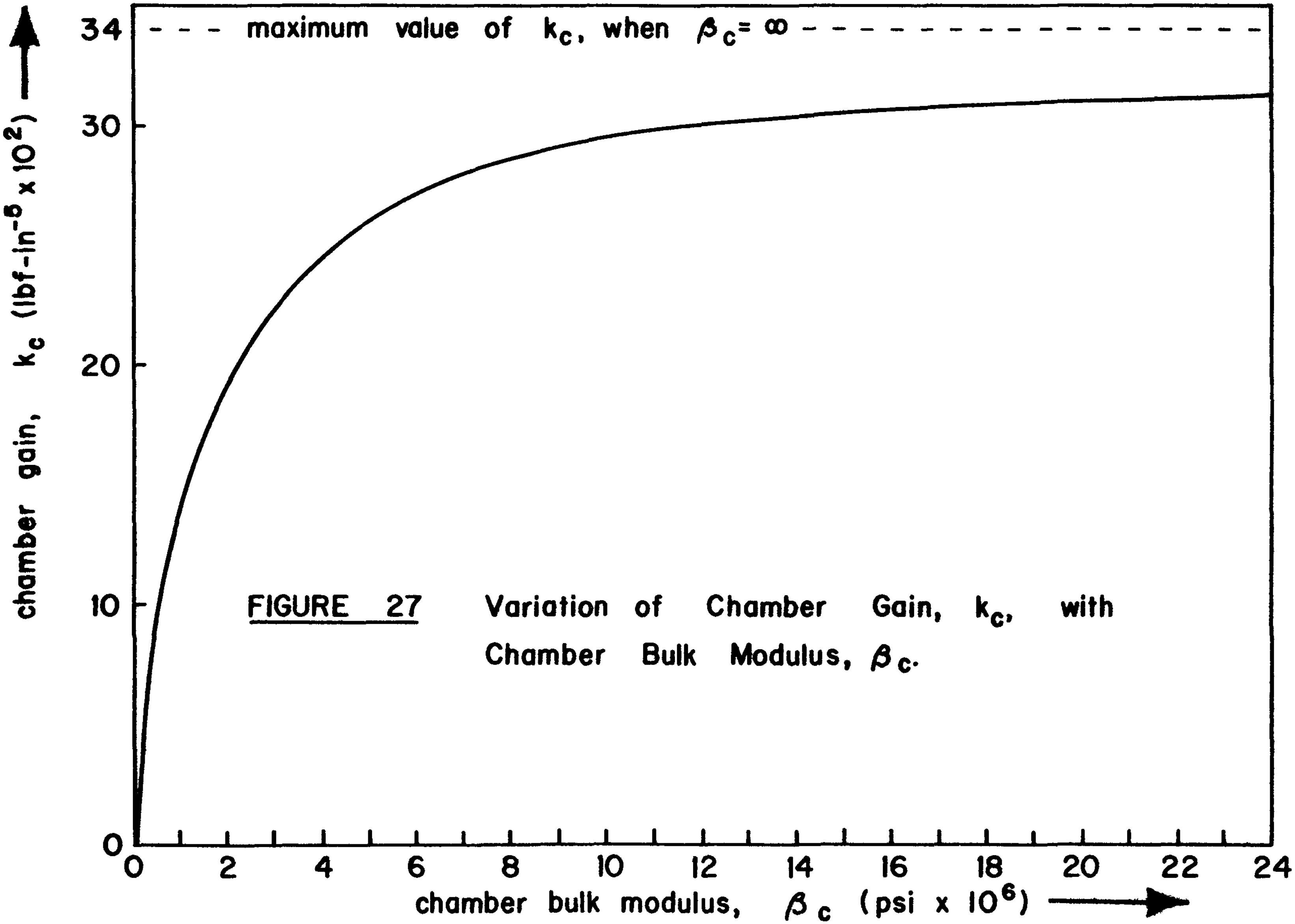


FIGURE 27 Variation of Chamber Gain, k_c, with Chamber Bulk Modulus, β_c.

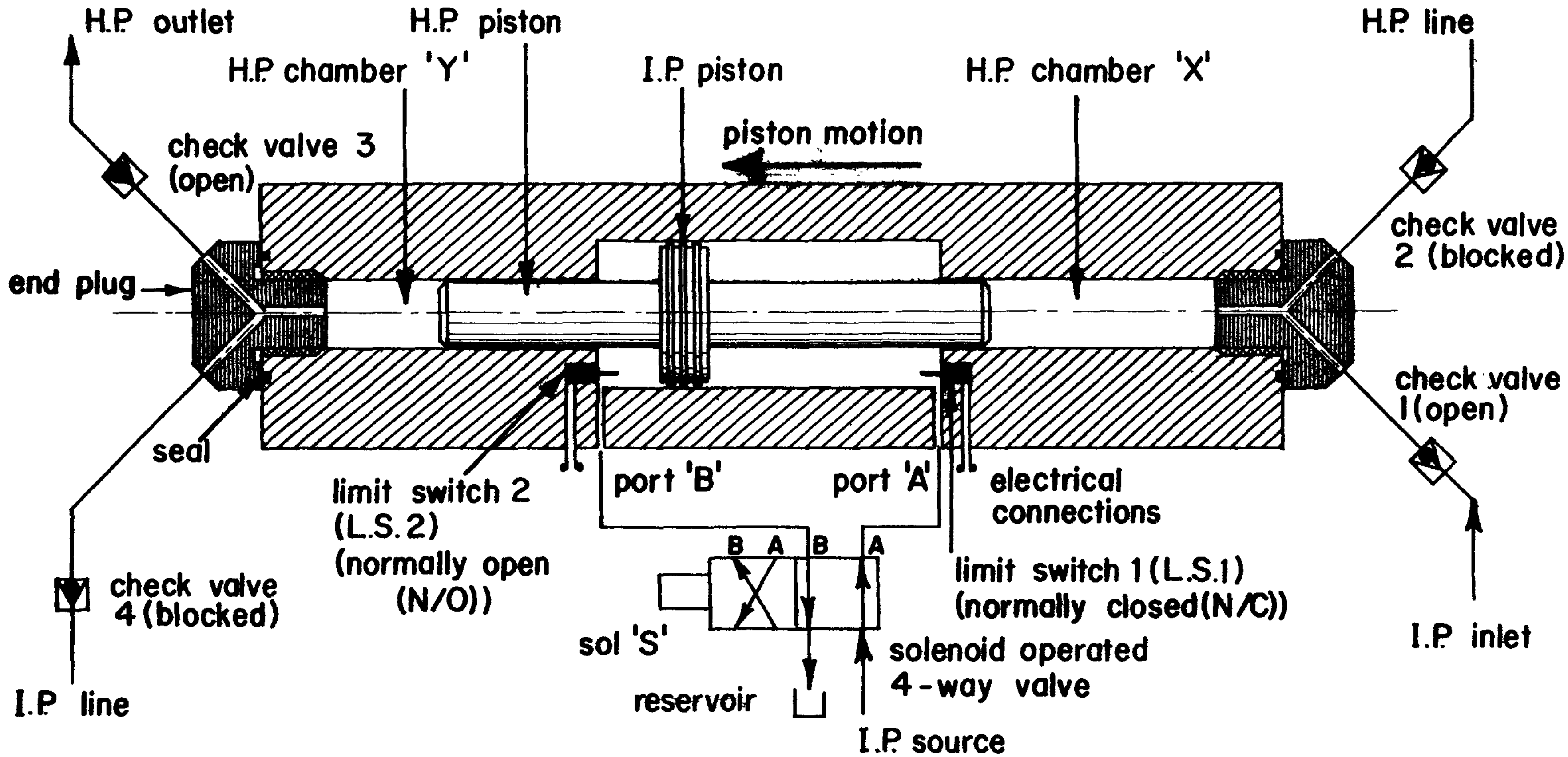


FIGURE 28(a) Schematic of Single Cylinder Double Acting H.P Intensifier with Solenoid Operated Reversing Mechanism.

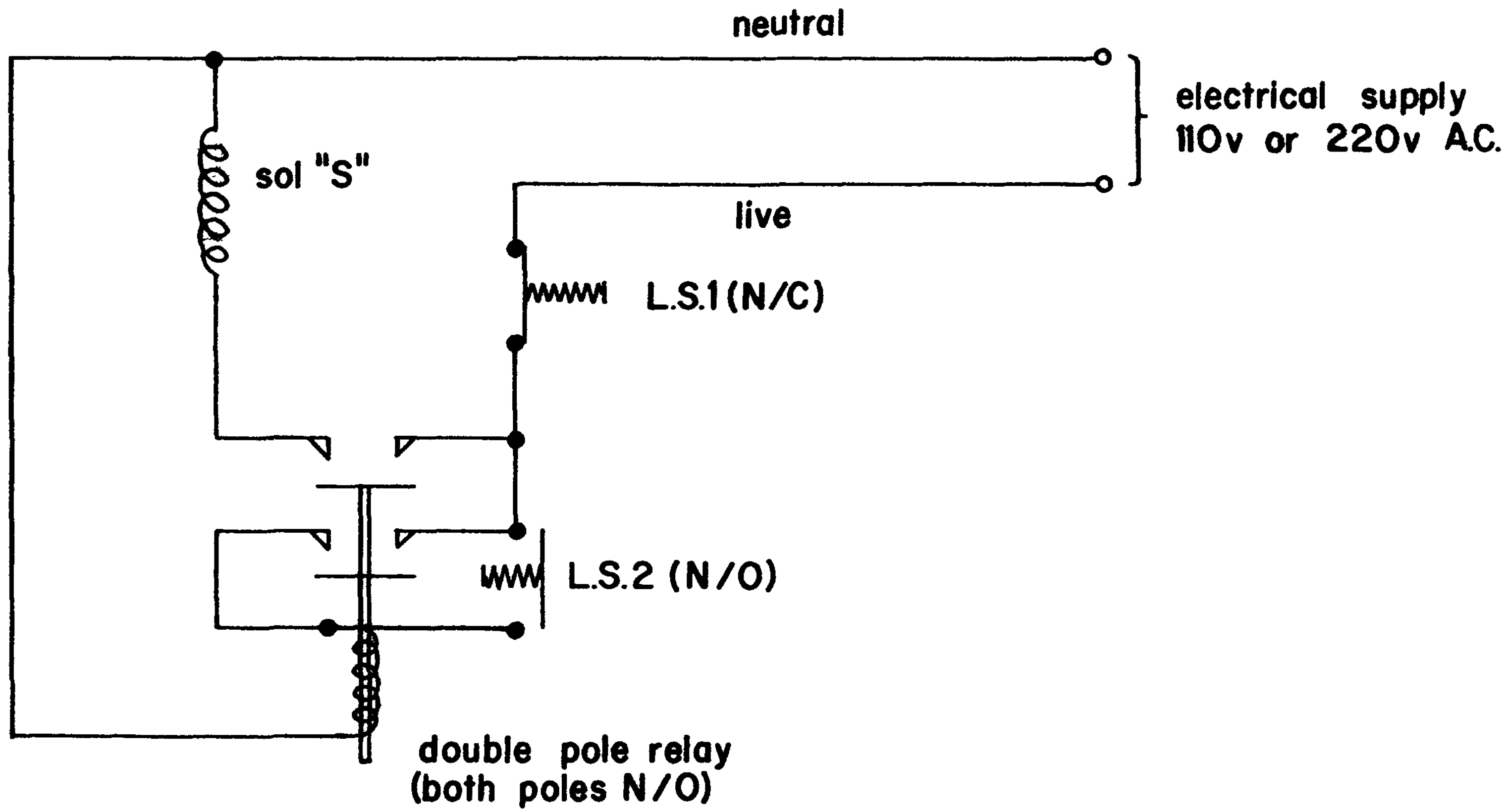


FIGURE 28(b) Electrical Reversing Circuit for H.P. Intensifier

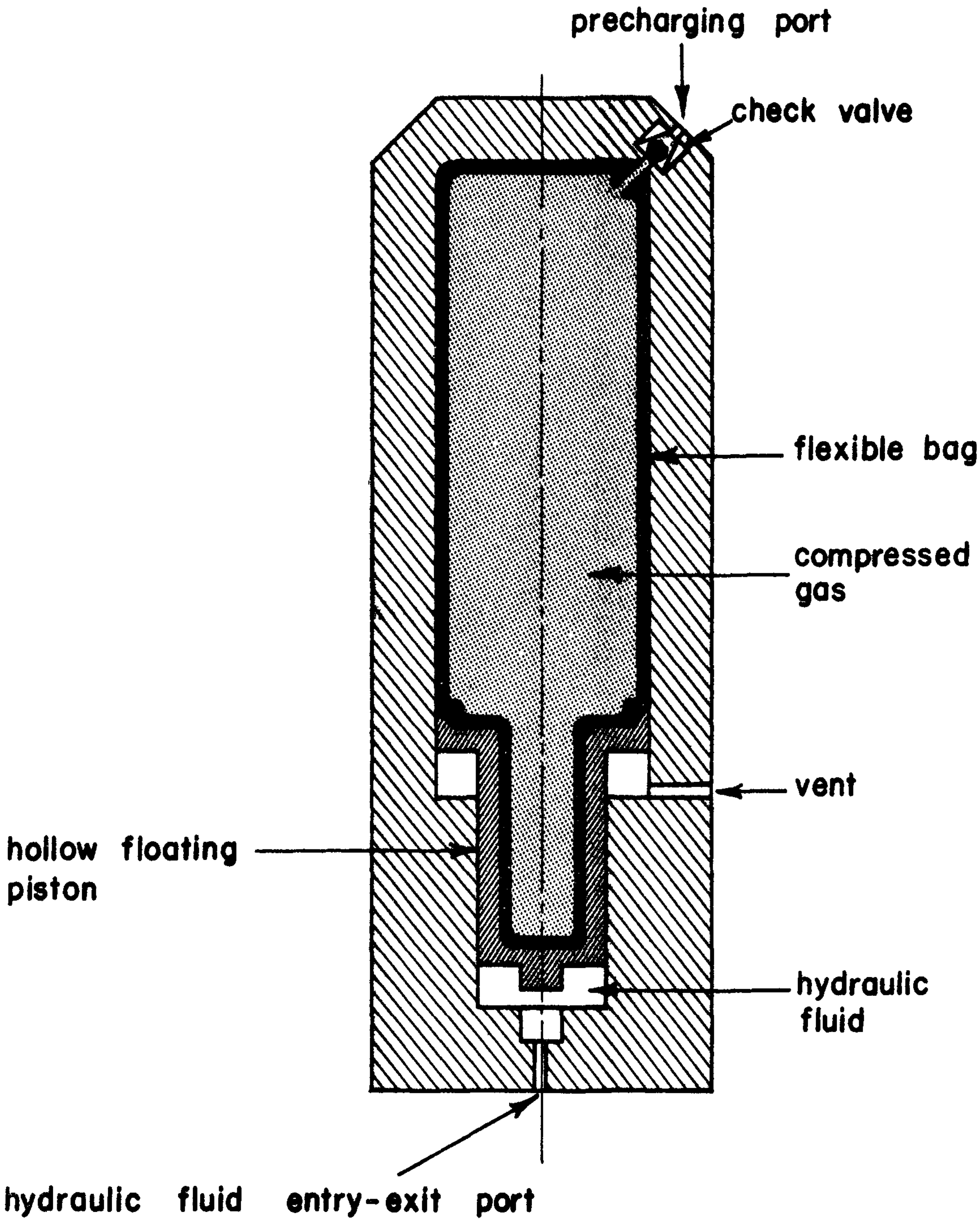


FIGURE 29 Compressed Gas Loaded H.P. Hydraulic Accumulator

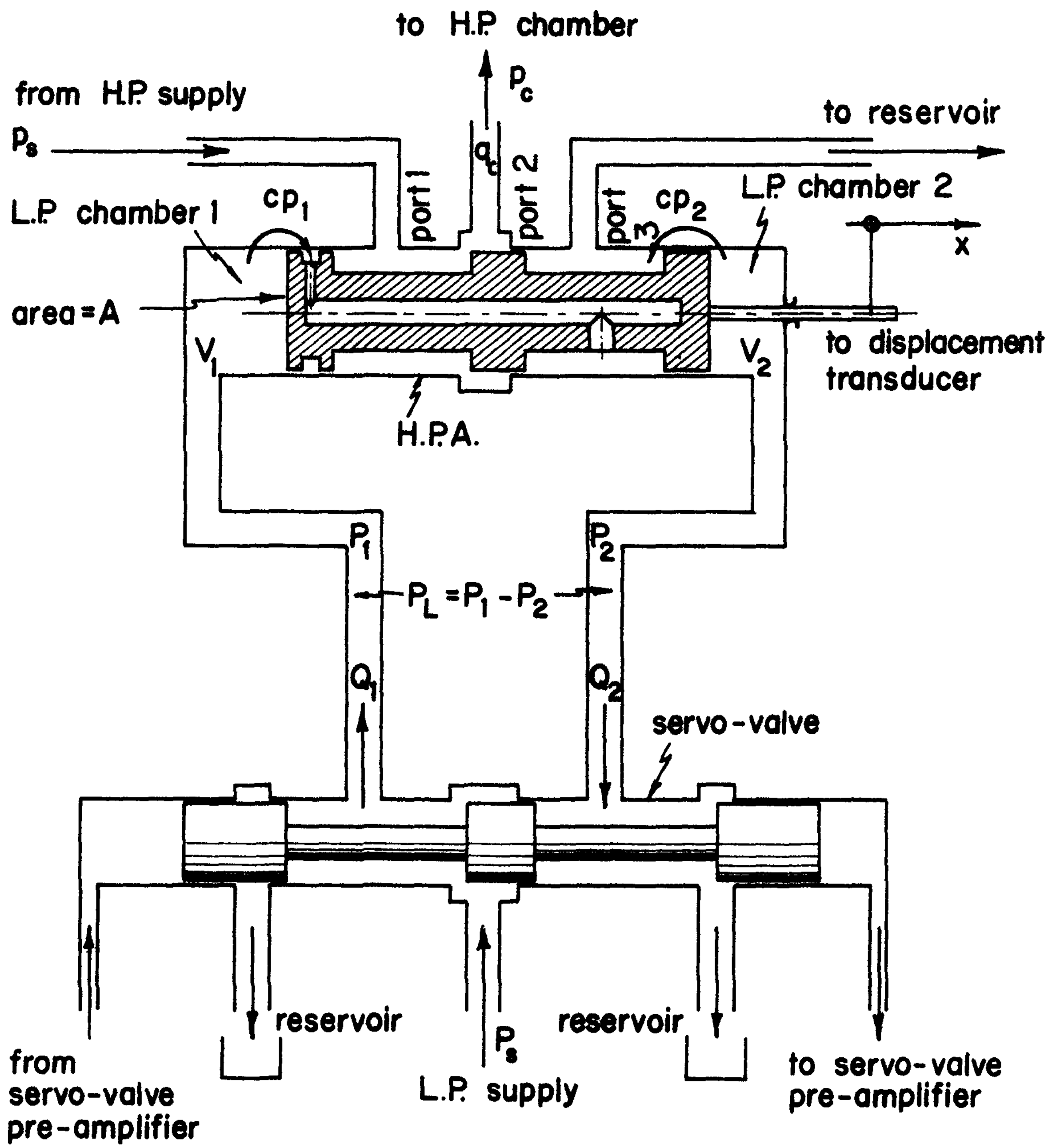
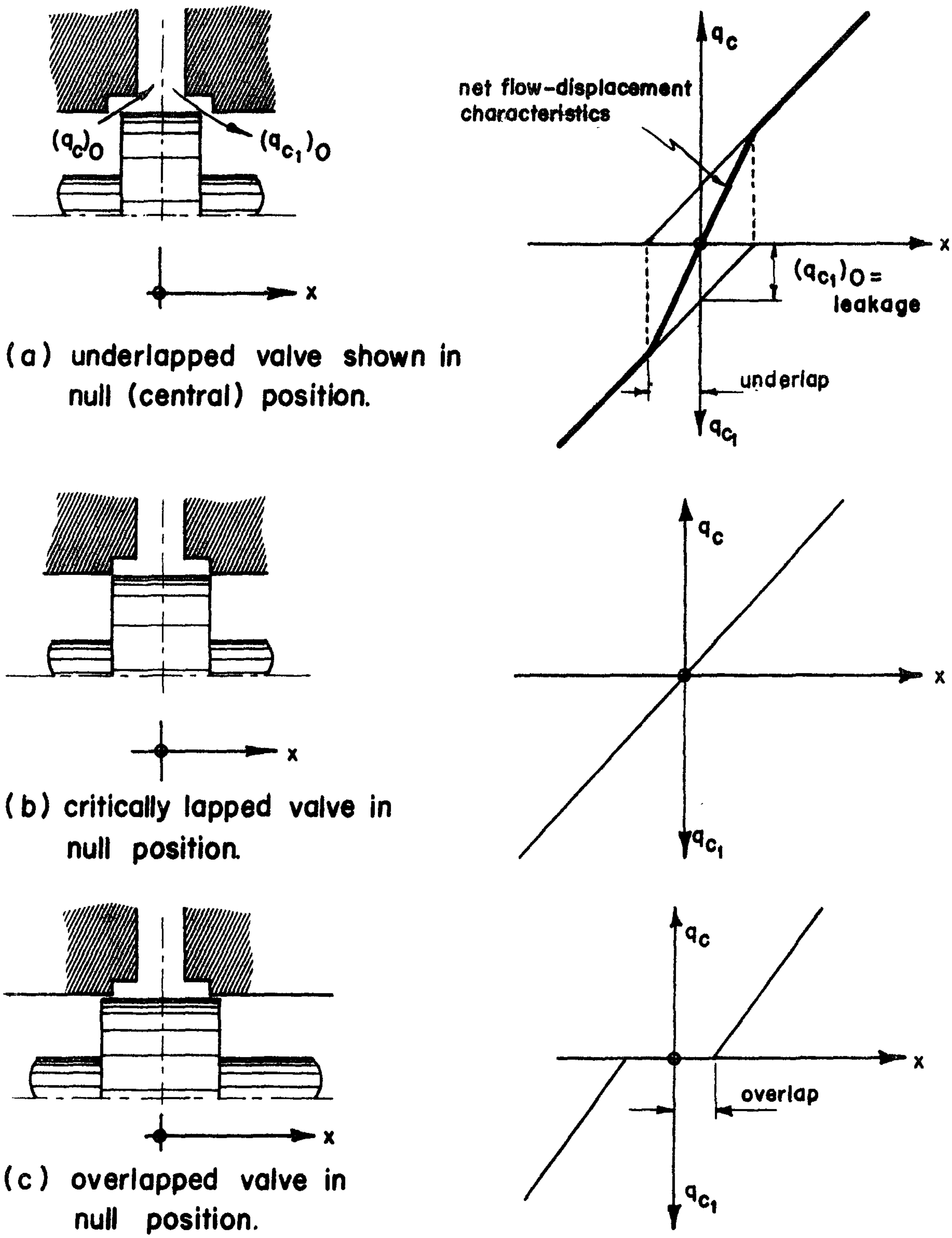


FIGURE 30 Schematic Representation of Servo-valve H.P.A. Operation.

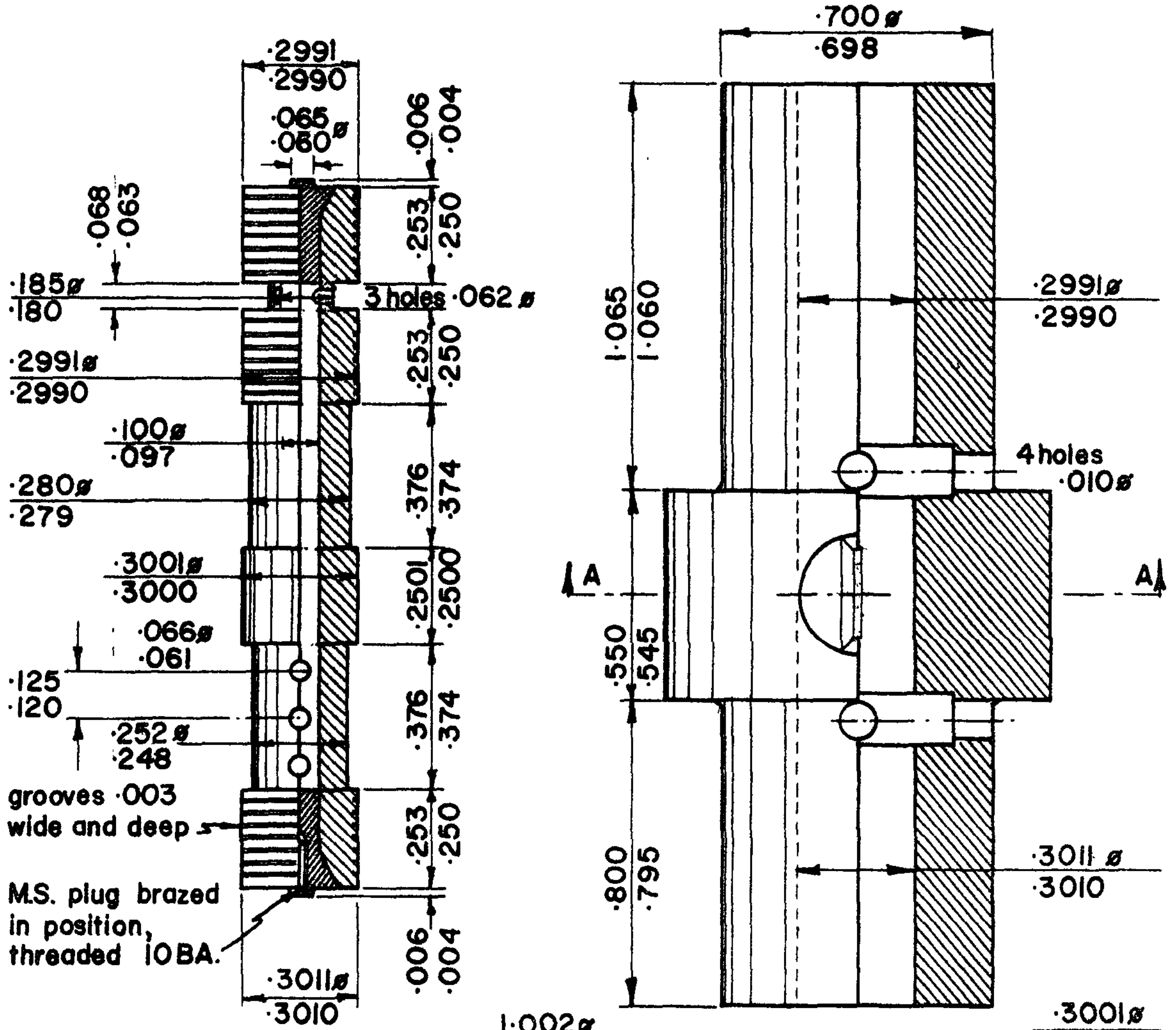


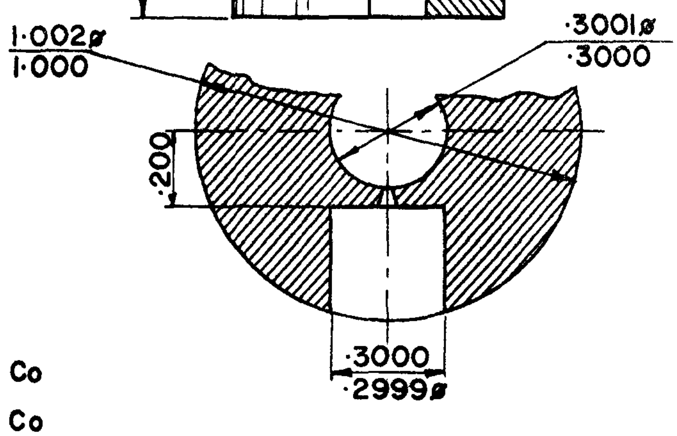
(a) underlapped valve shown in null (central) position.

(b) critically lapped valve in null position.

(c) overlapped valve in null position.

FIGURE 31 Ideal Flow-displacement Characteristics of a Valve with a Rectangular Orifice, Connected to the H.P. Chamber.





SECTION A A

NOTE i) all dimensions in inches

ii) both orifice slits are $\begin{matrix}0{\cdot}0050\\0{\cdot}0049\end{matrix}$ wide

and $\begin{matrix}0{\cdot}2500\\0{\cdot}2499\end{matrix}$ long

iii) material -

spool - tung. carb. 15 % Co

liner - " " 5 % Co

FIGURE 32 Leading Dimensions of Spool and Cylinder Liner Before Lapping.

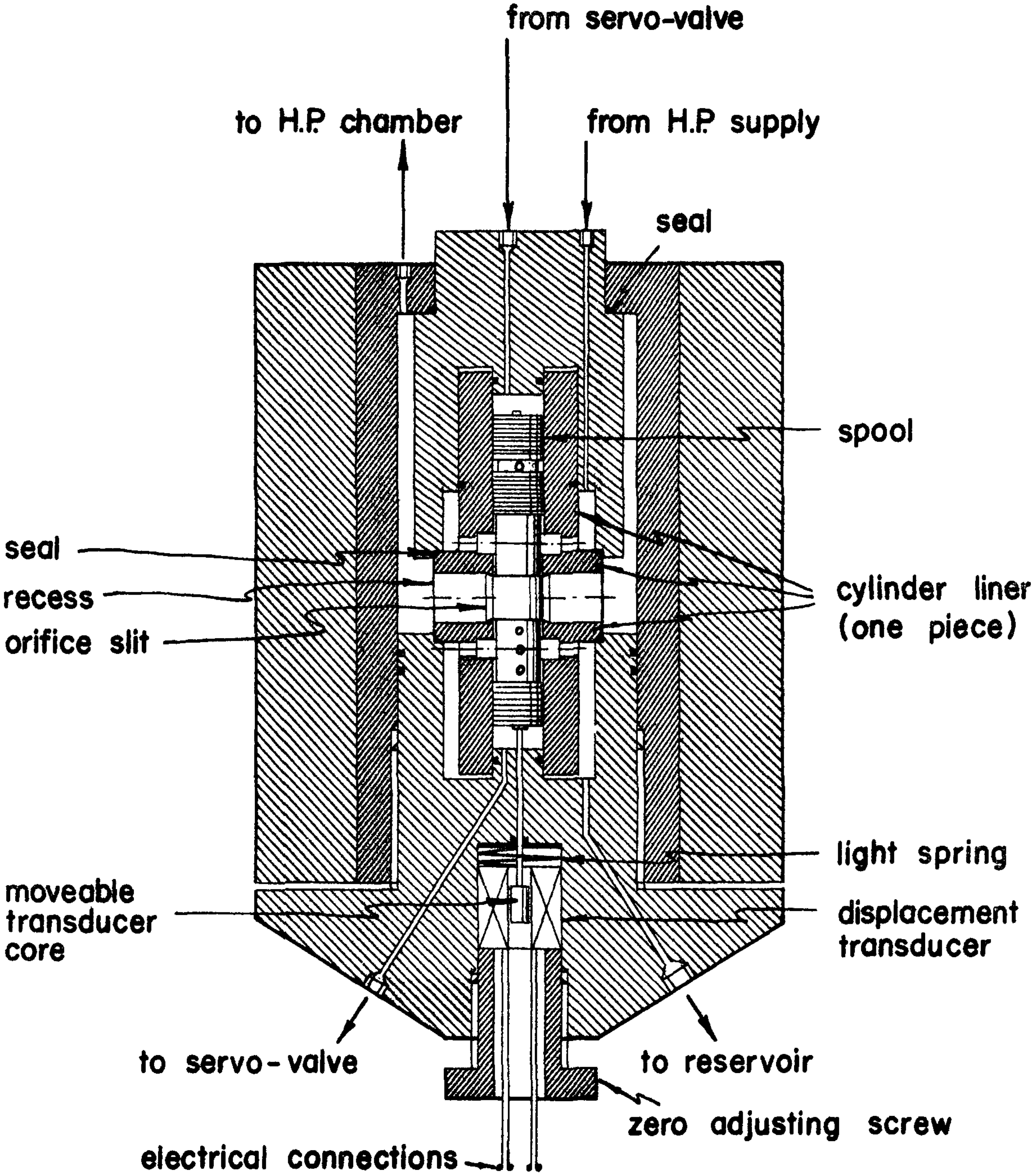


FIGURE 33 HPA Assembly.

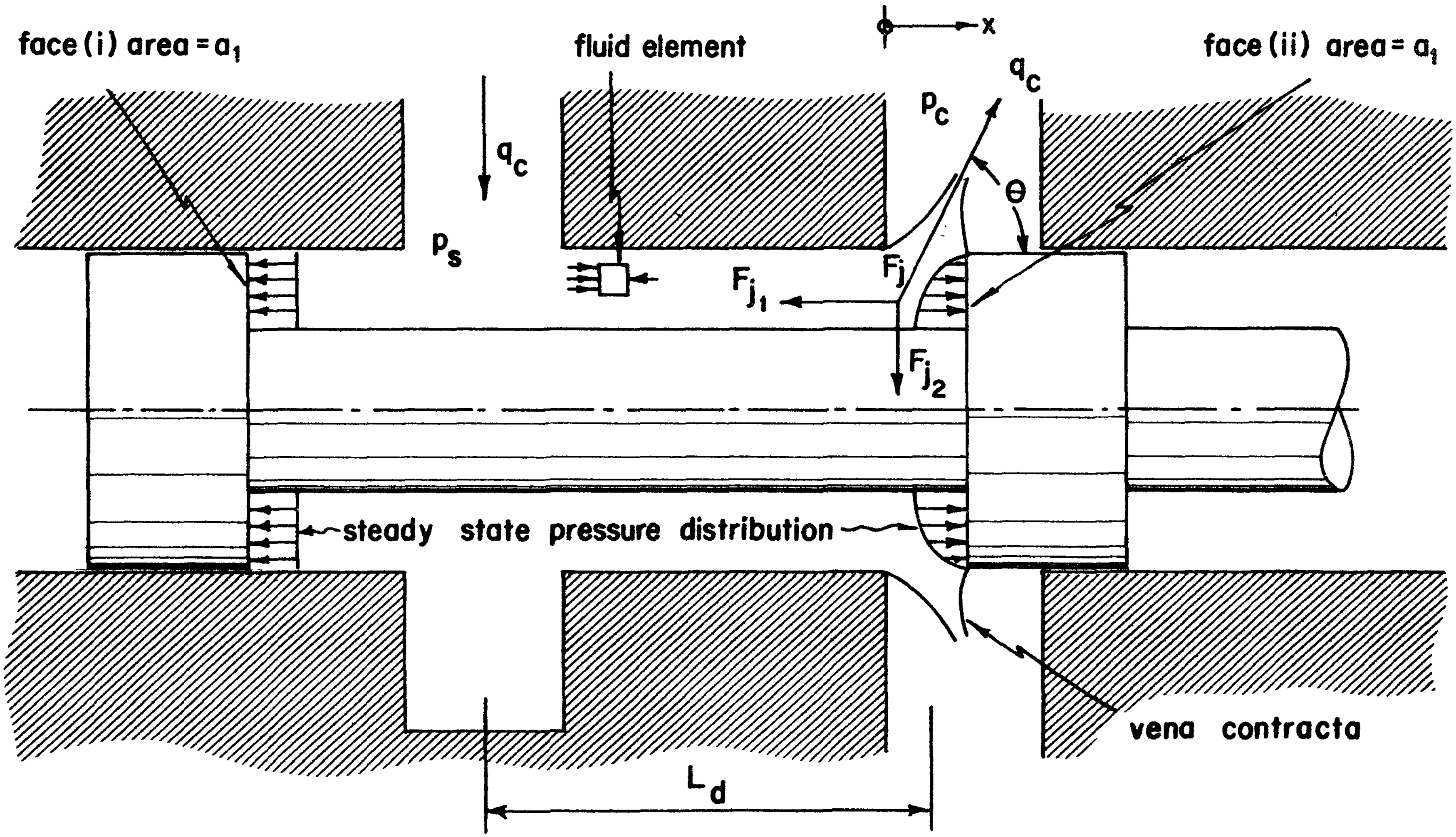


FIGURE 34 Flow Forces on a Spool Valve.

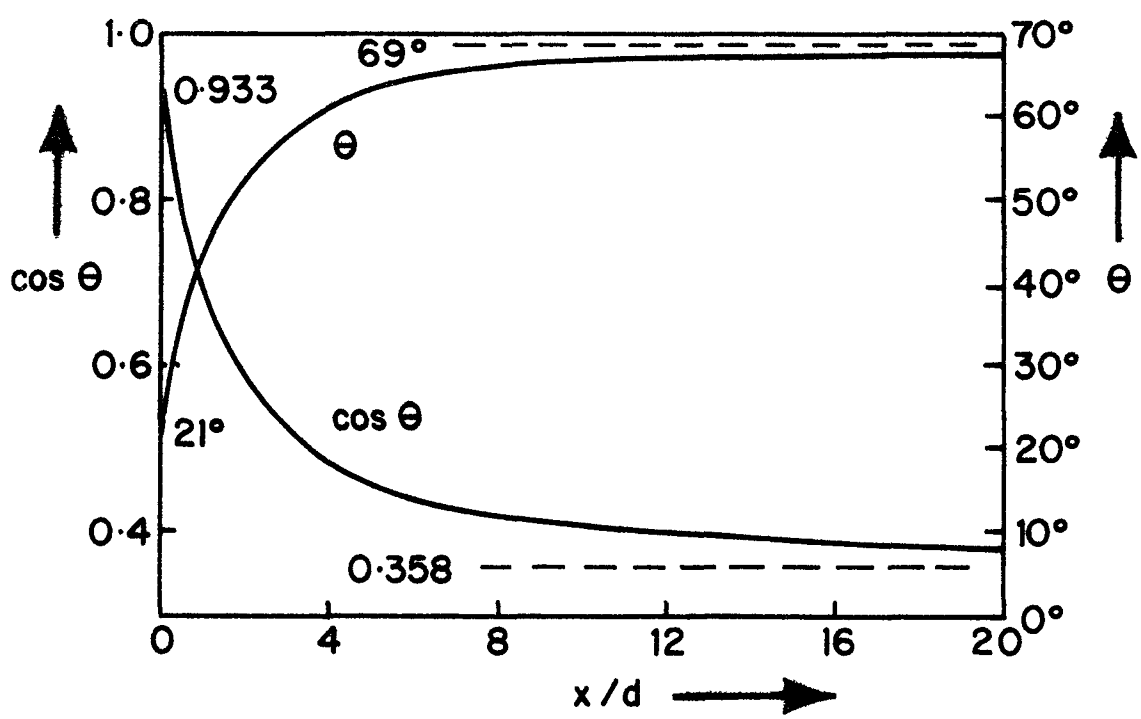


FIGURE 35 Variation of Jet Angle and its Cosine with the Ratio of Valve Opening to Radial Clearance.

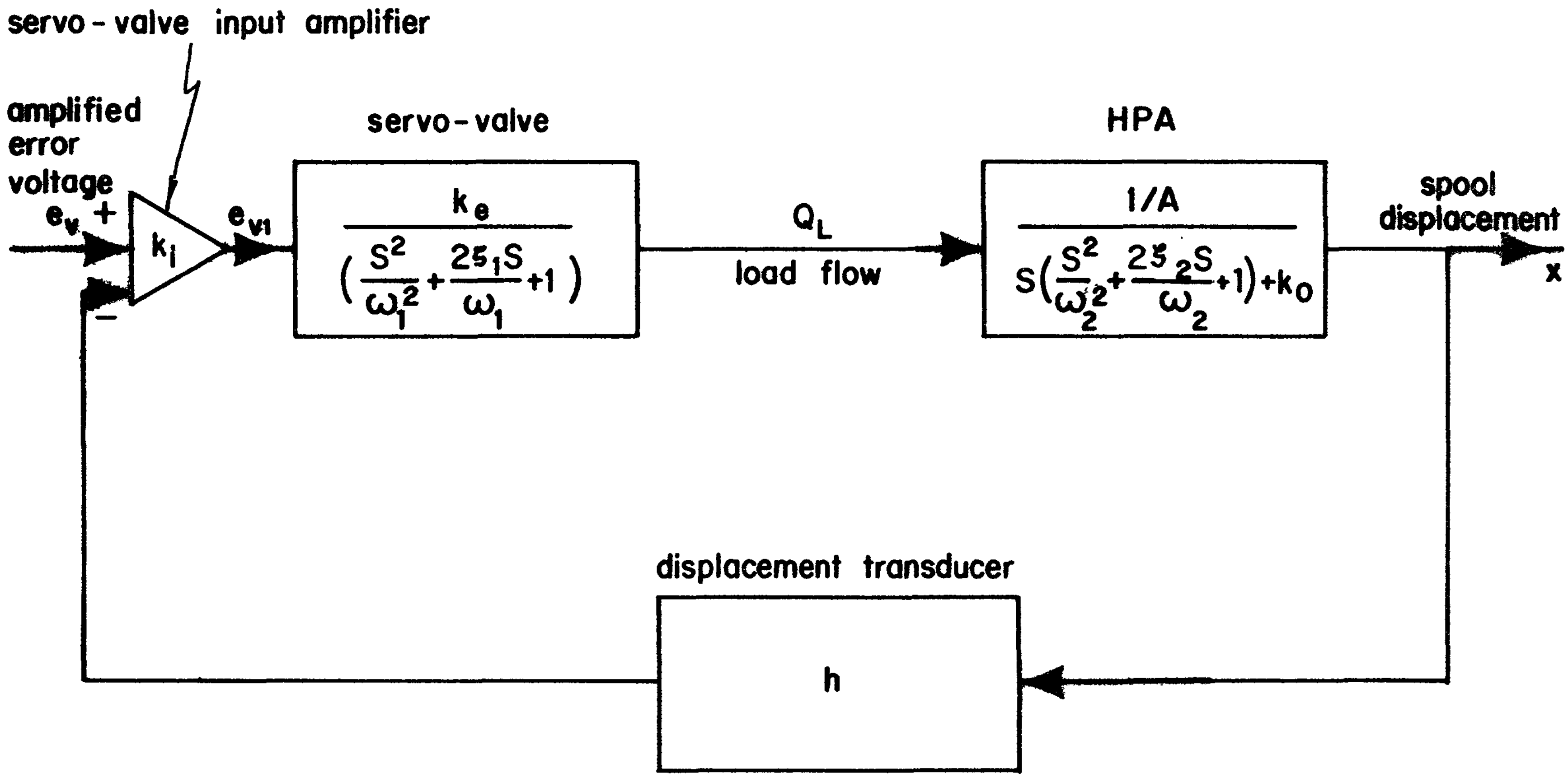
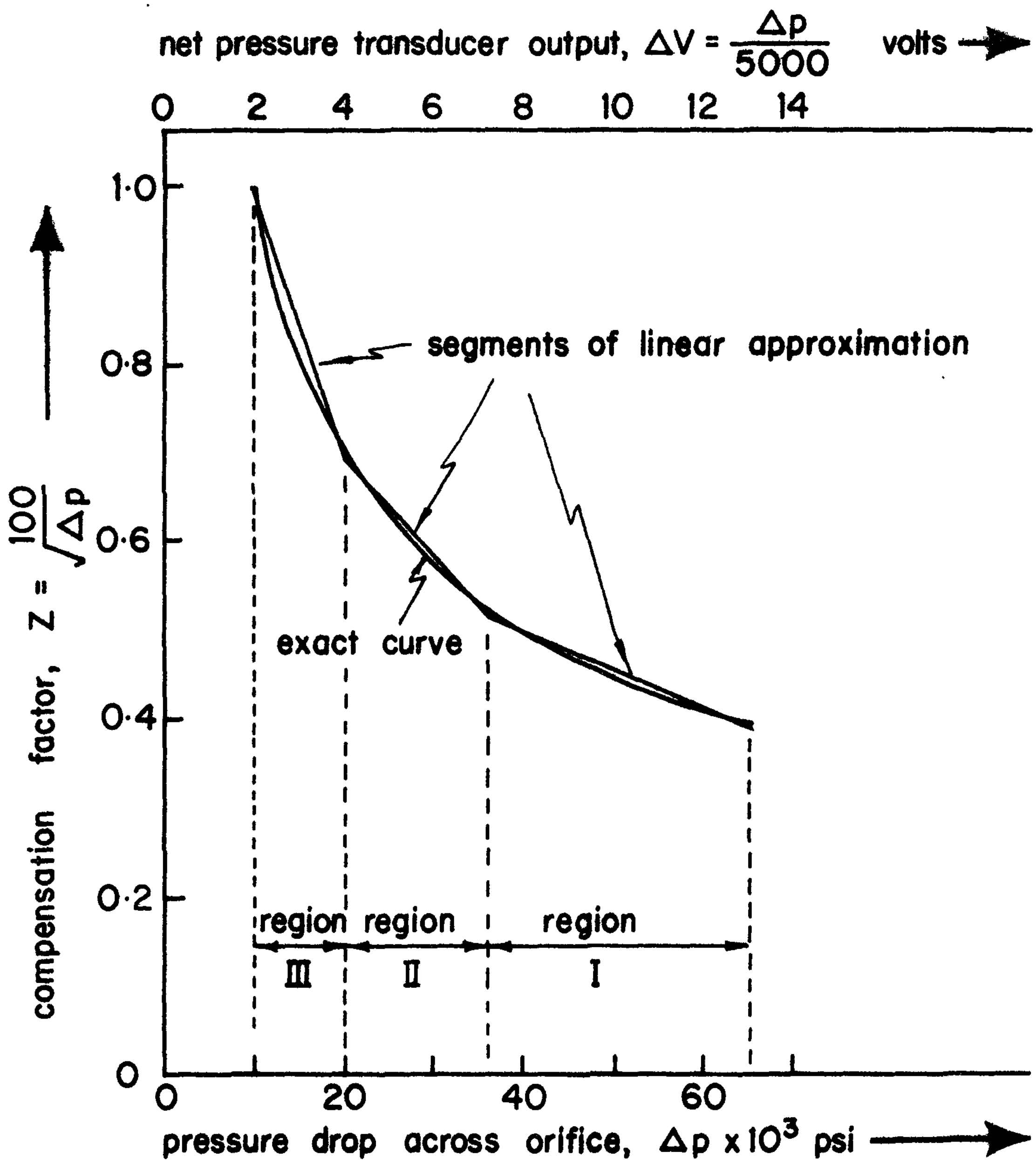


FIGURE 36 The Minor Loop of the Control System.



<u>note</u> for $x > 0$, $\Delta p = p_s - p_c$
for $x < 0$, $\Delta p = p_c$

<u>FIGURE 37</u> Compensation Factor as a Function of Orifice Pressure Drop, and its Linear Approximation.

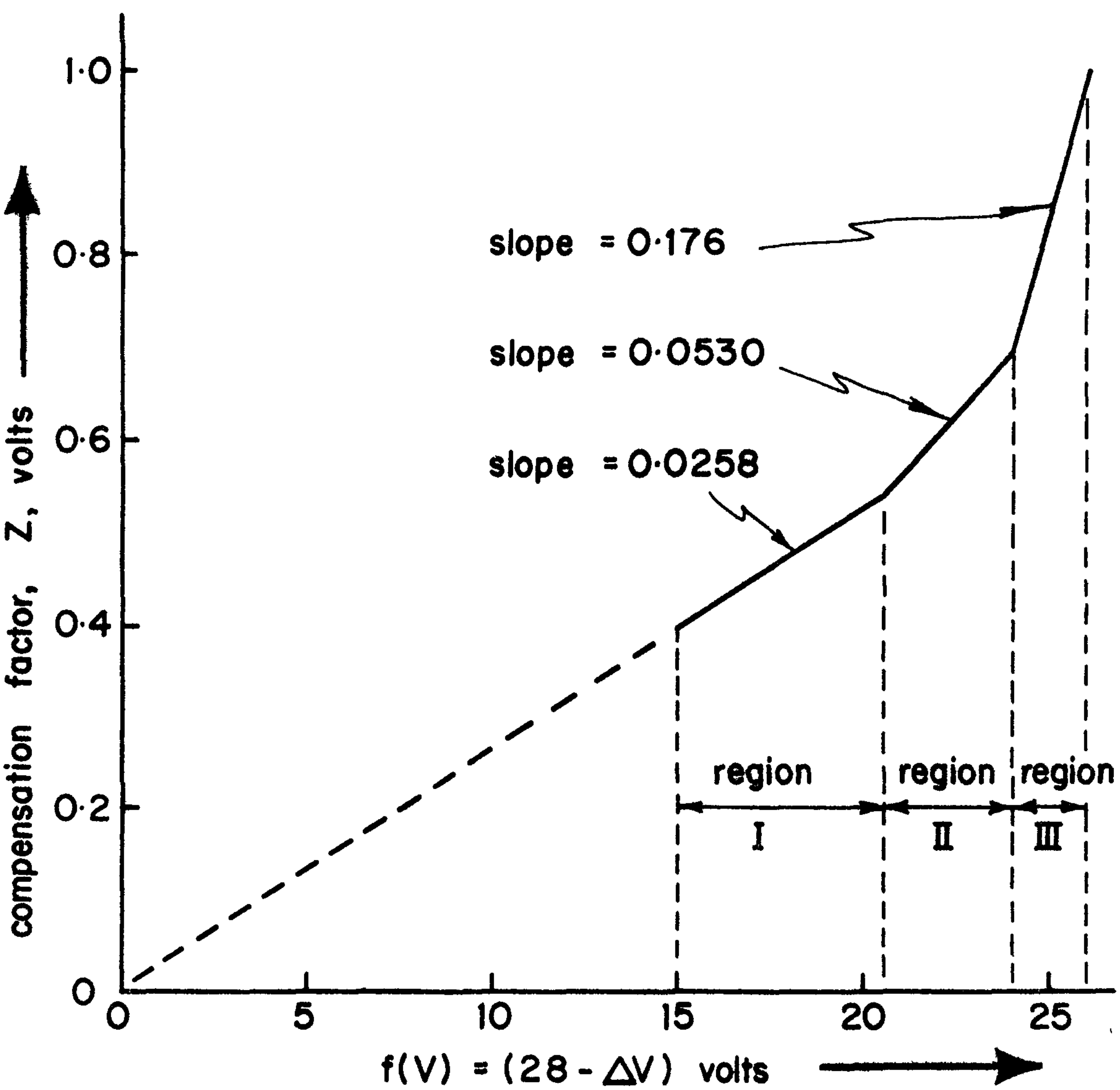


FIGURE 38 Linearised Variation of Gain Compensation Factor with Voltage Function, f(V).

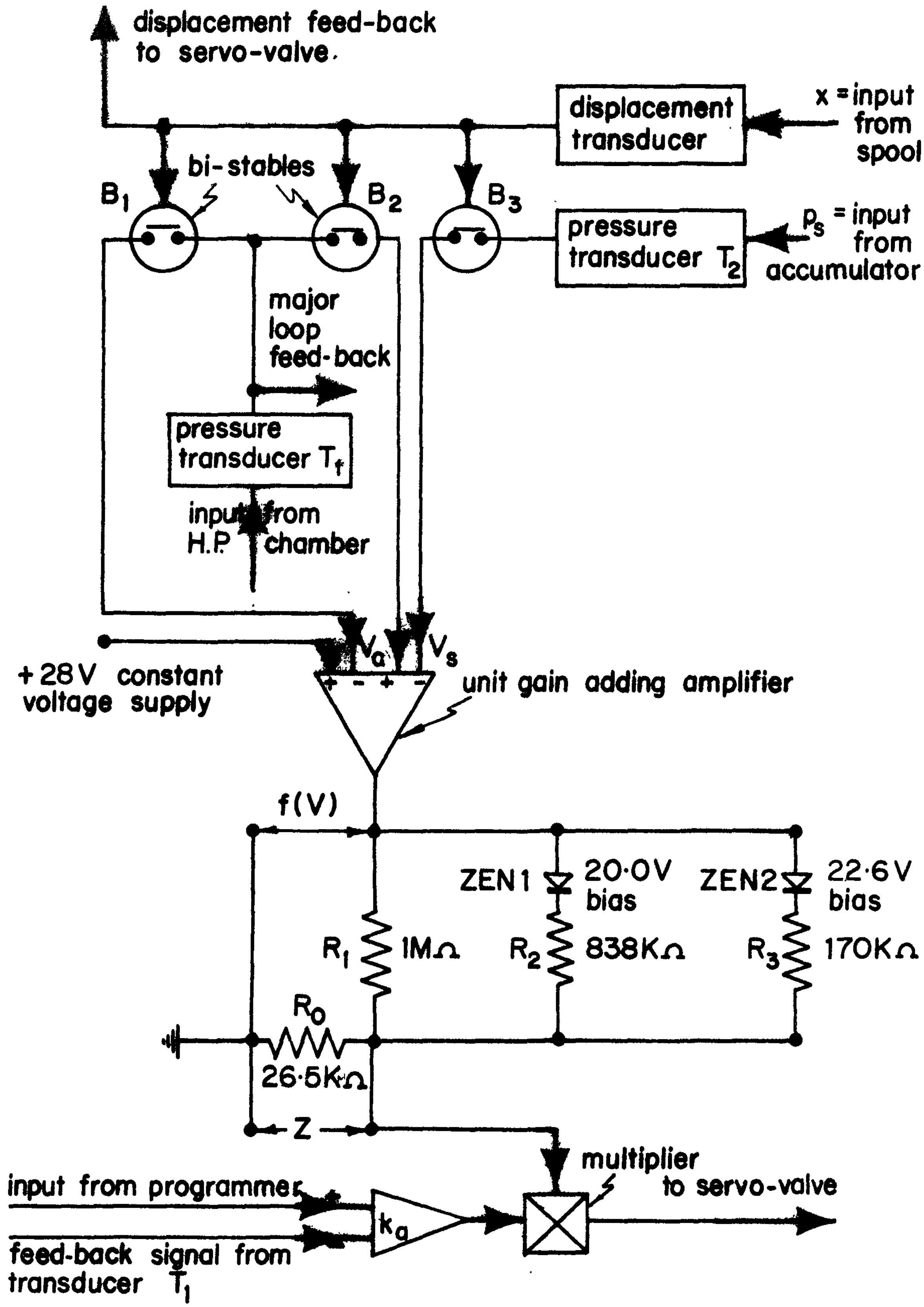


FIGURE 39 Adaptive Gain Compensator.

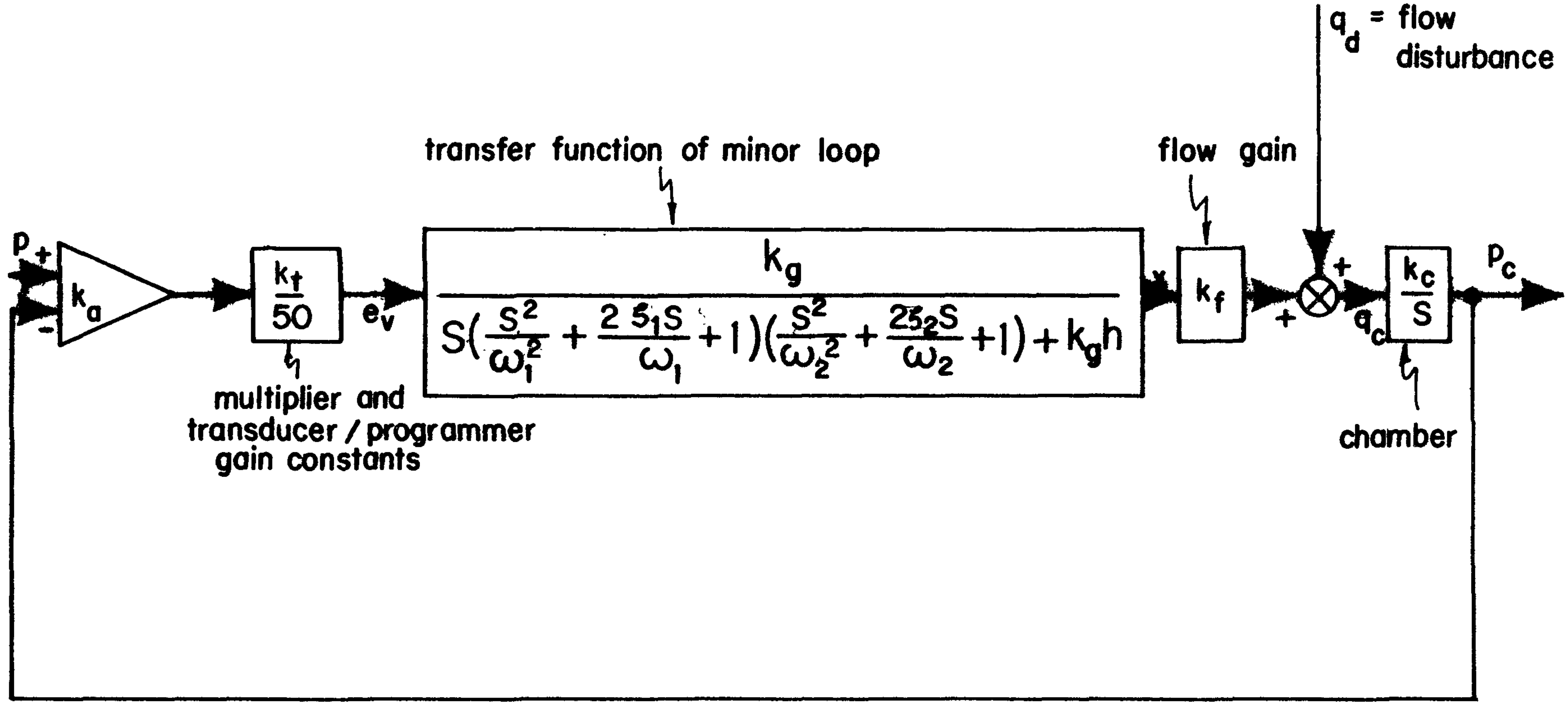


FIGURE 40 Major Loop of Control System.

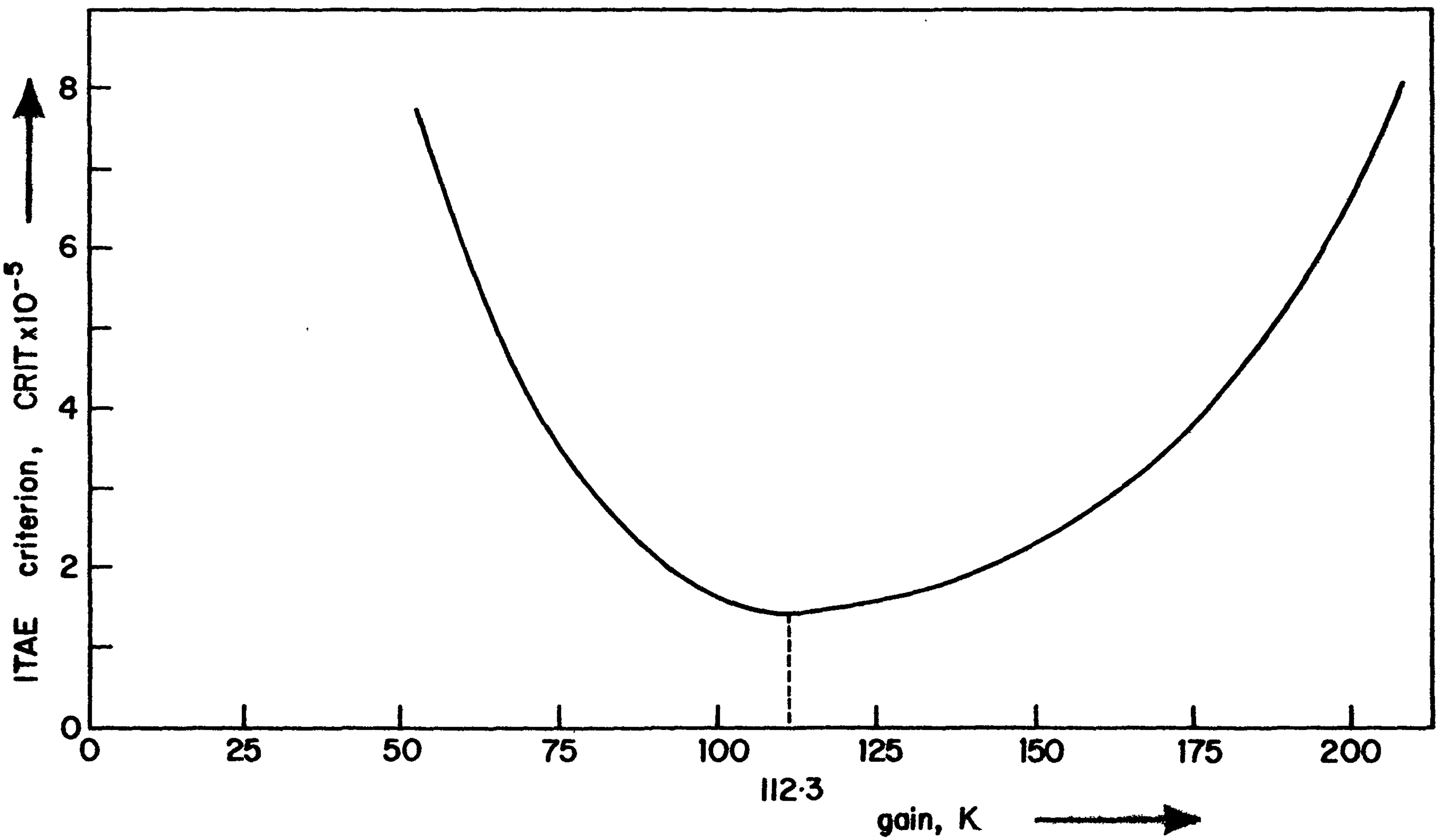


FIGURE 41 Sensitivity of the System Performance to Variations in the Gain, K.

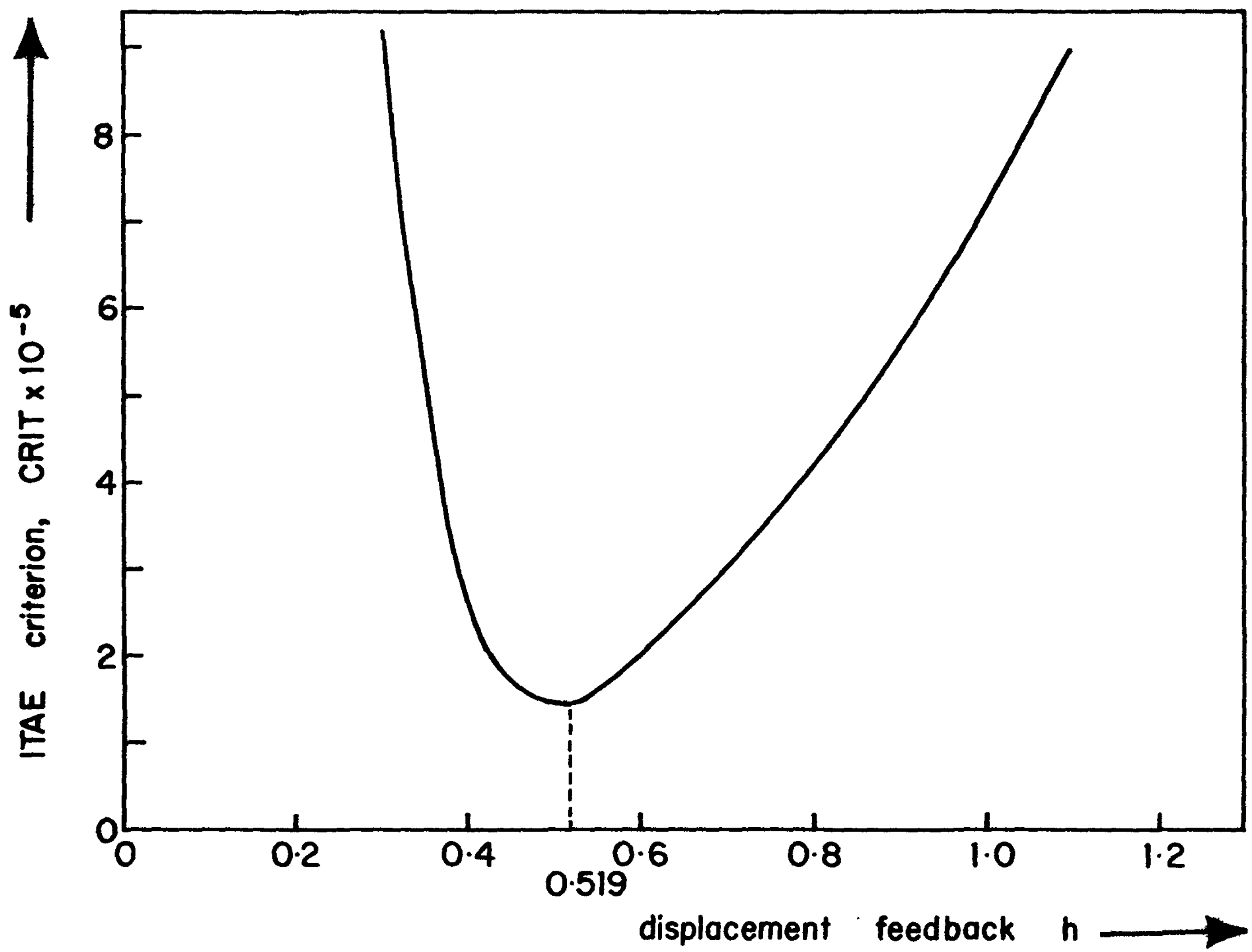


FIGURE 42 Sensitivity of the System Performance to Variations in the Displacement Feedback h.

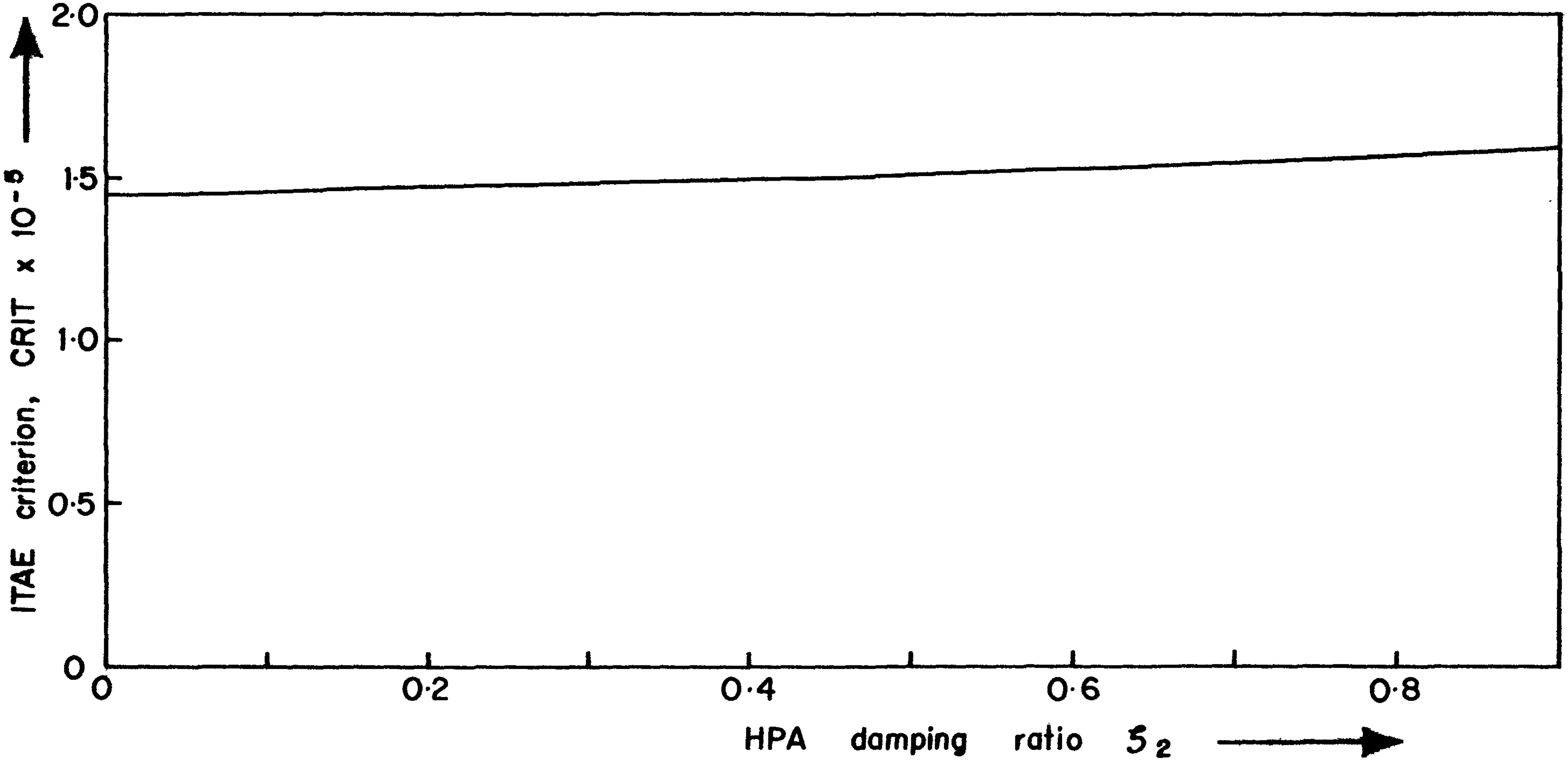


FIGURE 43 Sensitivity of the System Performance to Variations in the HPA Damping Ratio, ζ_2.

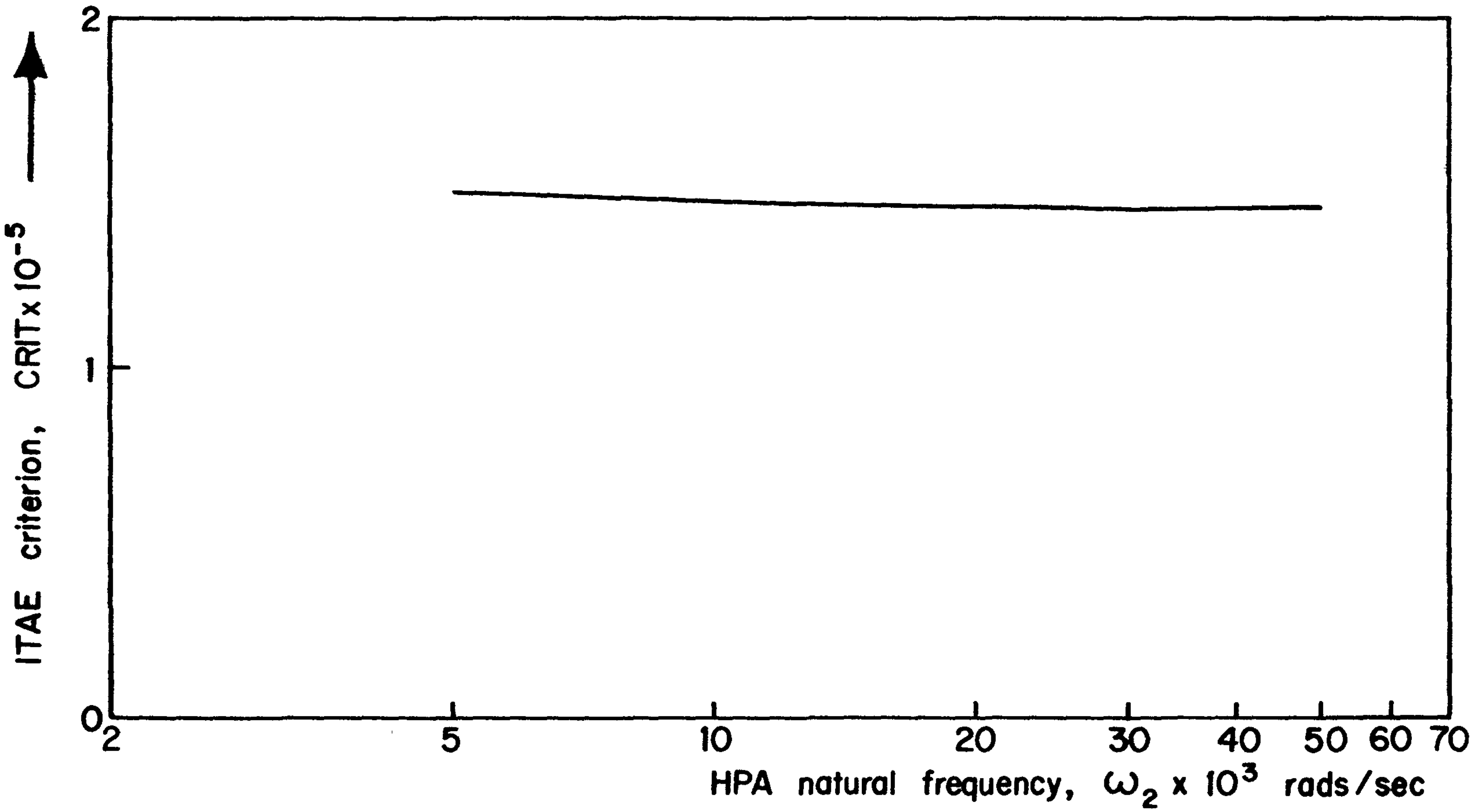


FIGURE 44 Sensitivity of the System Performance to Variations in the HPA Natural Frequency, ω_2.

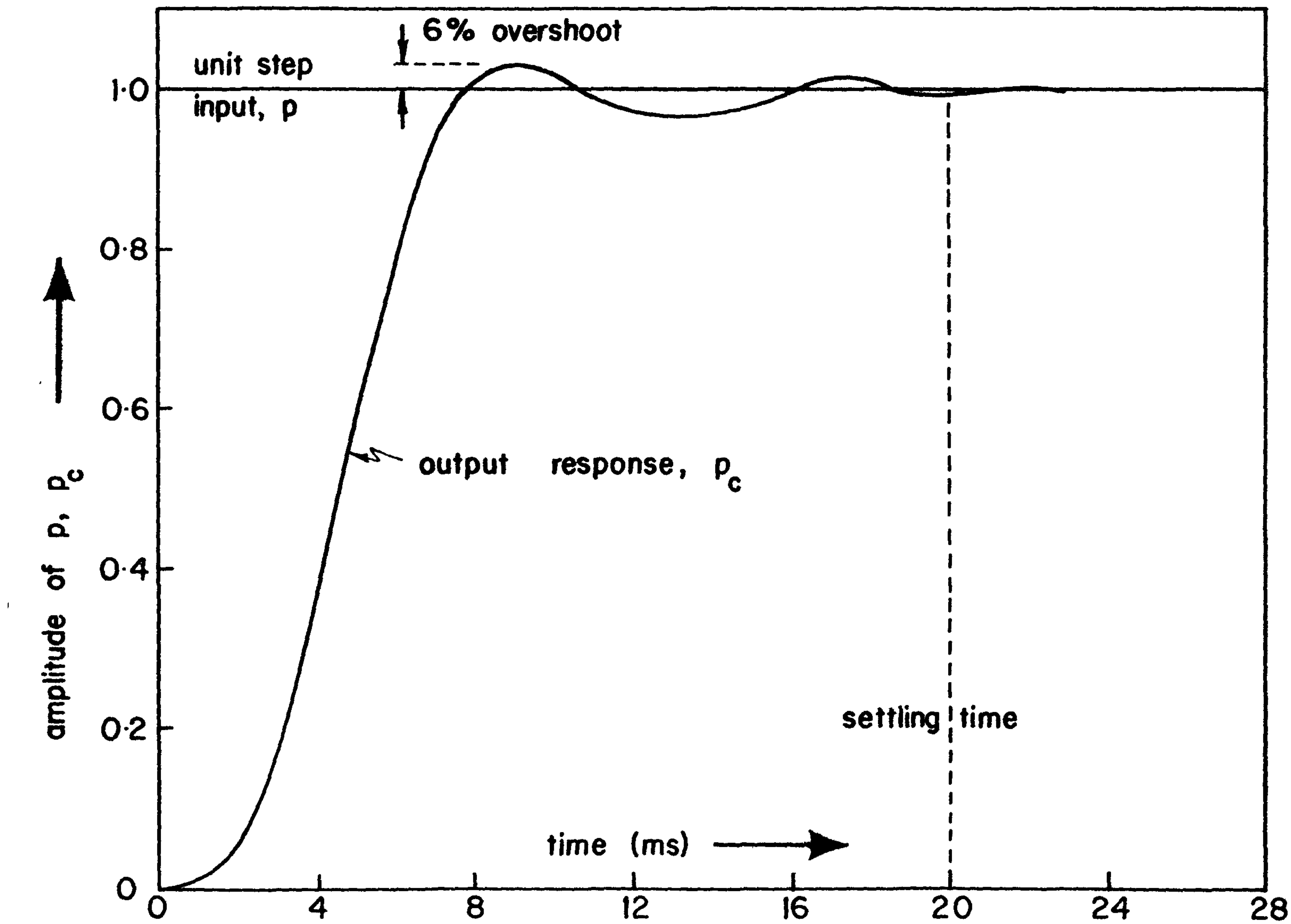


FIGURE 45 Transient Response of Output to a Unit Step Input for the Optimised System.

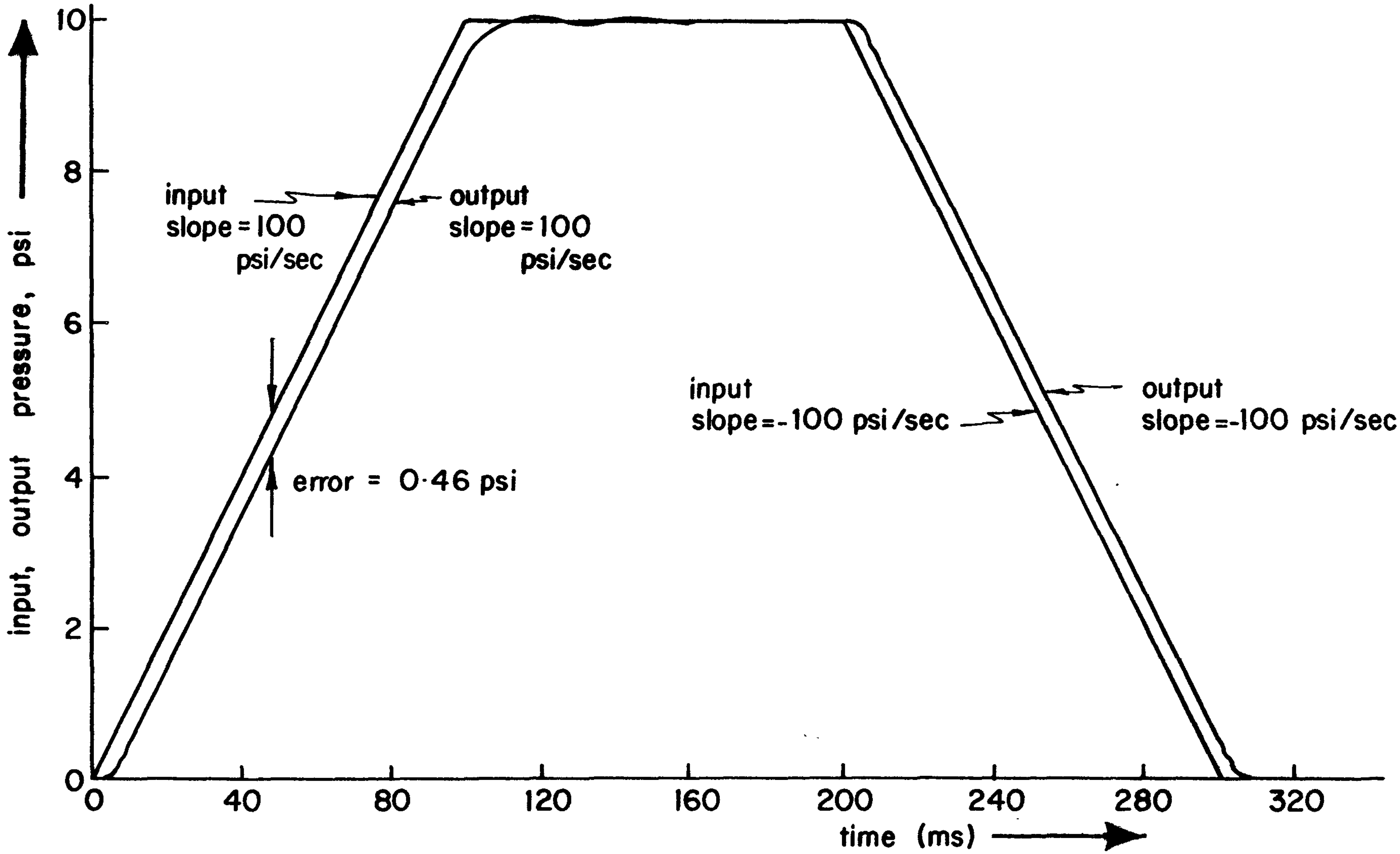


FIGURE 46 Simulation of Response of Optimised System to a 'Plateau' Input.

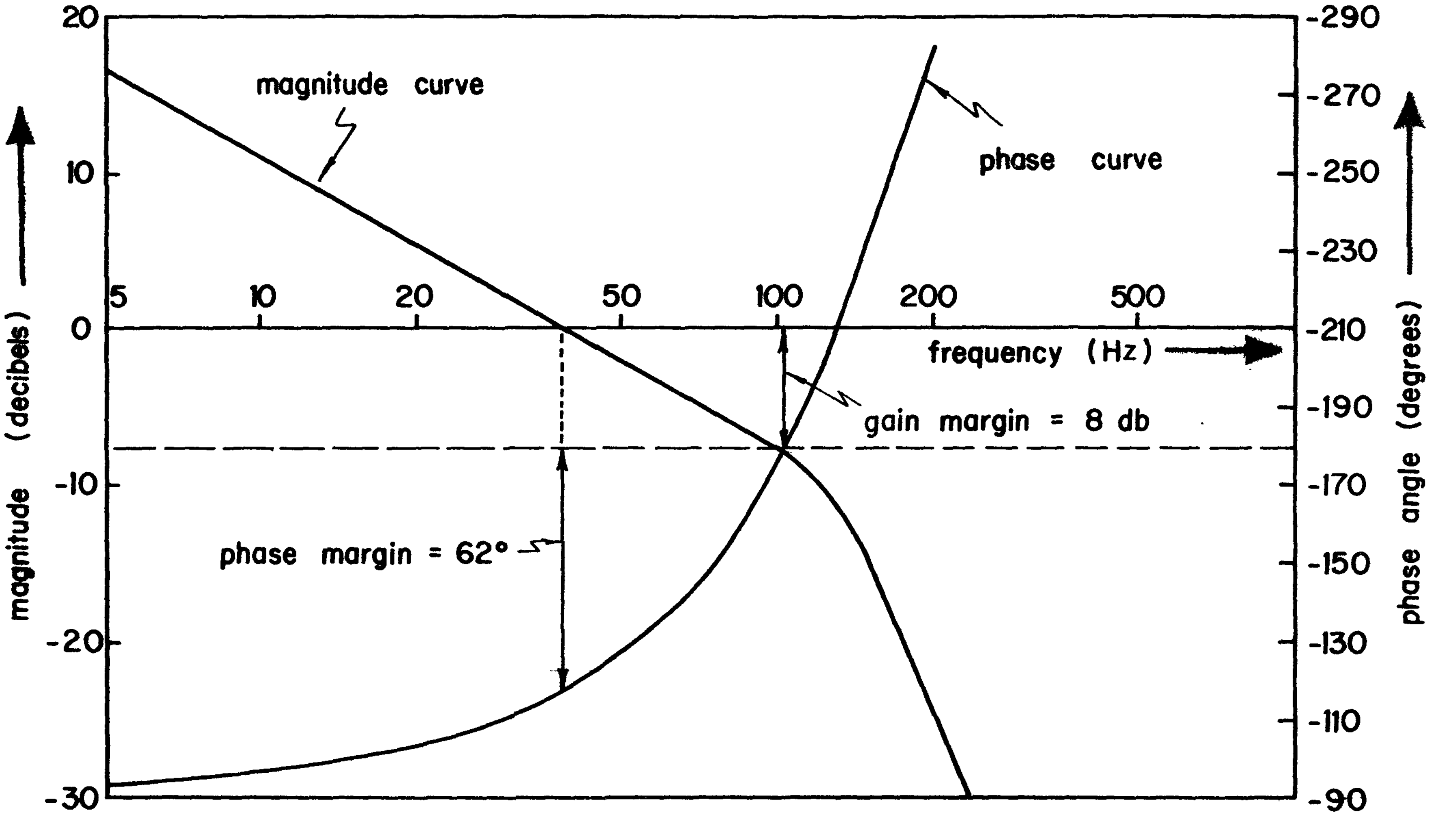


FIGURE 47 Open Loop Frequency Response of Optimised System.

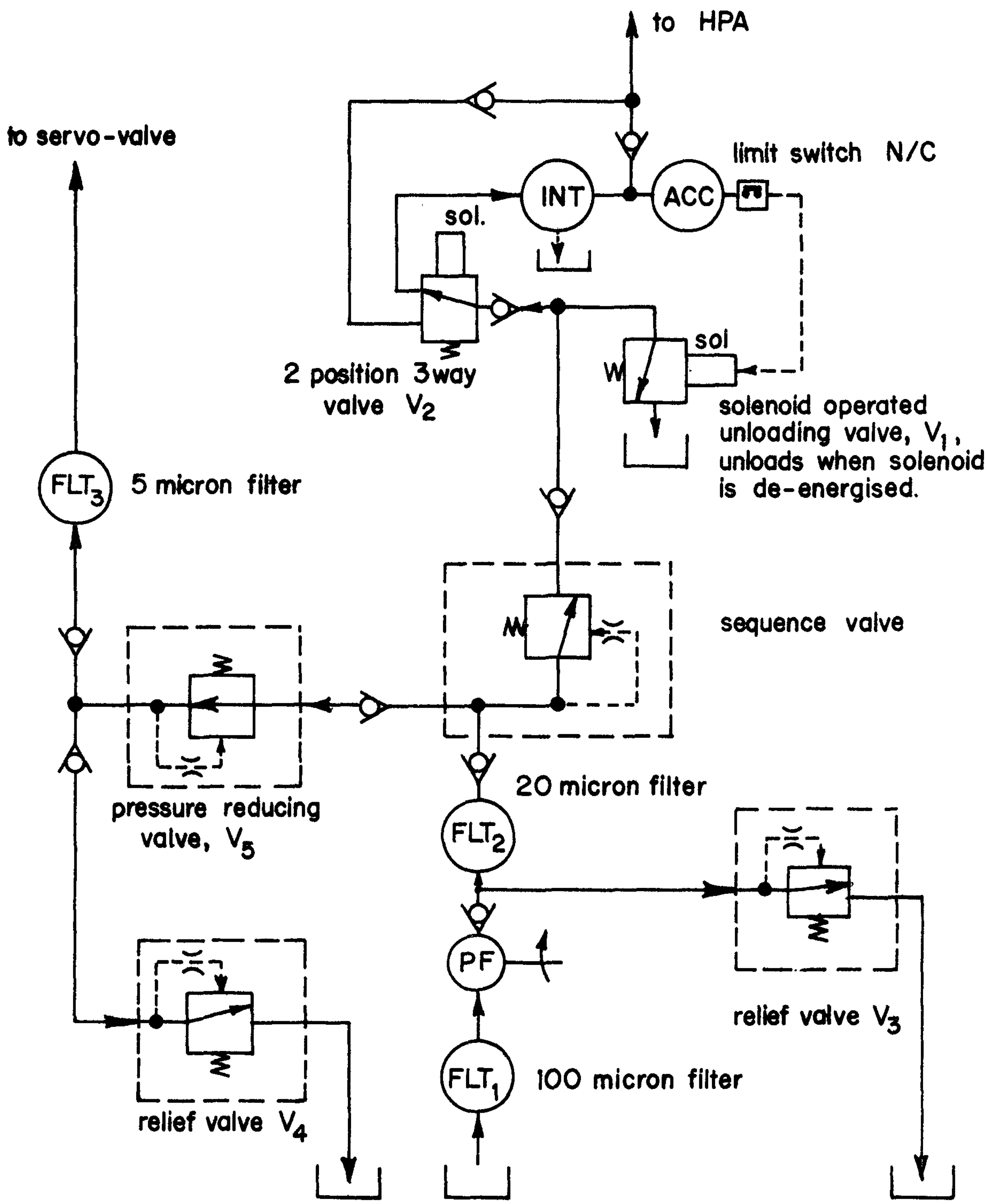
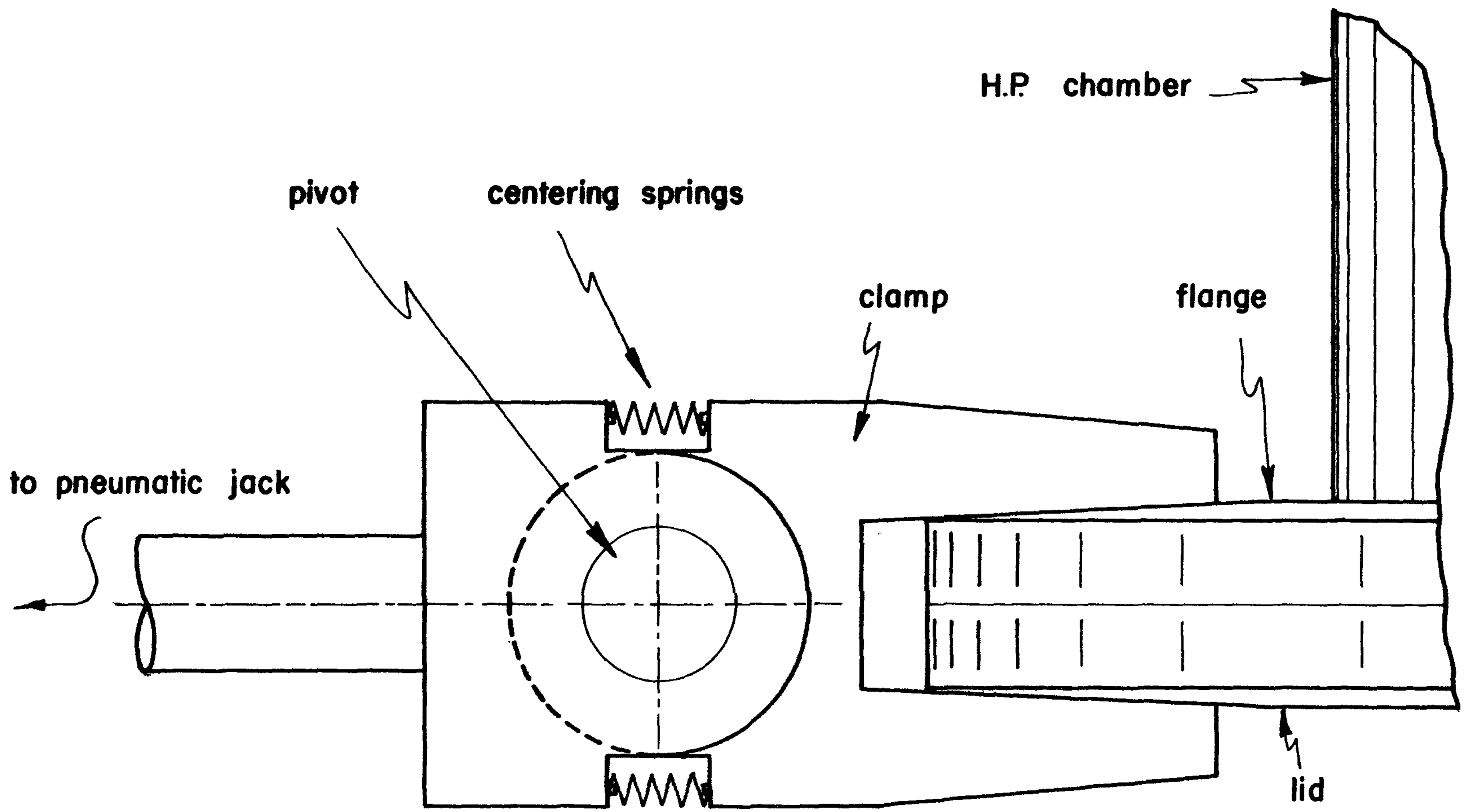


FIGURE 48 Practical Schematic of Hydraulic Power Supply System.



<u>FIGURE 49</u> Clamp for H.P. Chamber Lid.

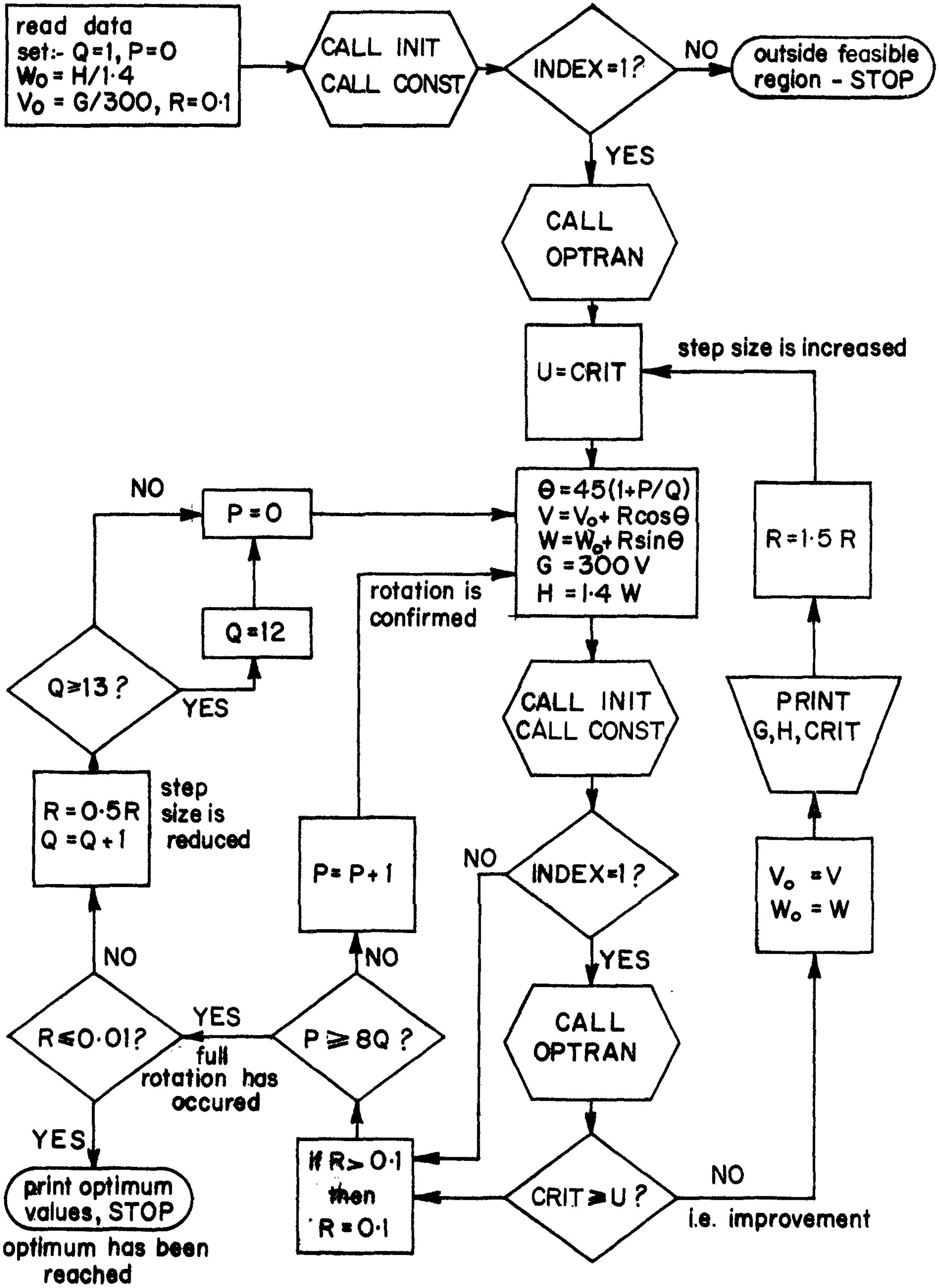


FIGURE 50 Flow Chart for Direct Search Optimisation.

REFERENCES

1. Meyfarth, H. W. (Ed.)
"Hydraulic 16,500 p.s.i. System for Fluid Forming Presses"
J. Assoc. Prof. Engng., ONT.
V 29, No. 10
Nov. 1968.

2. Thomas, S. L. S., Turner, H. S., and Wall, W. F.
"Piston-Cylinder High Pressure Apparatus for Use up to 25 Kilobars"
High Pressure Eng. Conf.
Inst. of Mech. Engrs.
London, Sept. 1967.

3. Rolfe, B. W.
"Cold Isostatic Presses: Equipment and Techniques used at A.W.R.E."
High Pressure Eng. Conf.
Inst. of Mech. Engrs.
London, Sept. 1967.

4. Skinner, J. L., Daniels, R. D. and Sliepcevich, C. M.
British Chemical Engng.
No. 8,
London, 1963.

5. Slater, H. K., and Green, O.
"Augmented Hydrostatic Extrusion of Continuous bar"
High Pressure Eng. Conf.
Inst. of Mech. Engrs.
London, Sept. 1967.

6. Lowe, A. H., Donaldson, C. J., and Wilkinson, P. T.
"Developments in Hydrostatic Extrusion at the National Engineering Laboratory"
High Pressure Eng. Conf.
London, Sept. 1967.

7. Crawley, J., Pennell, J. A., and Saunders, A.
"Some Problems in the Design and Development of Hydrostatic Extrusion Systems"
High Pressure Eng. Conf.
London, Sept. 1967.

8. Pugh, H. L. D.
J. Mech. Engng. Sci.
Vol. 6, 362.
1964.

9. Bridgman, P. W.
"Studies in Large Plastic Flow and Fracture"
Harvard Univ. Press
1964.

10. Minorsky, N.
"Non-Linear Oscillations"
Van Nostrand,
N.Y.
1964.

11. Manning, W. R. D.
"The Overstrain of Tubes by Internal Pressure"
Engineering, V. 159, 101 and 183.
London, 1945.

12. Franklin, G. J., and Morrison, J. L. M.
"Autofrettage of Cylinders"
Proc. Inst. Mech. Engrs., V. 174, p. 947
London, 1960.

13. Manning, W. R. D.
"The Design and Construction of Containers"
Bulleid Memorial Lectures, High Pressure Engineering. V. 2, No. 2.
Univ. of Nottingham, 1963.

14. Wilson, W. R. D. and Skelton, W. J.
"Design of High Pressure Cylinders"
High Pressure Eng. Conf.
Inst. of Mech. Engrs.
London, Sept. 1967.

15. Thiel, N. W.
"Diamonds and Some Industrial Applications"
S.A. Mech. Engr.
V. 19, p. 33
Johannesburg, 1969.

16. Bialy, L.
"Analysis of Stresses in Pressure Magnifying Anvils"
University of Witwatersrand
Johannesburg, 1966.

17. O'Donovan, K.
"A 2000 Ton Cubic Press for Synthetic Diamond Manufacture"
H.P.M.L. Report DD68.
Johannesburg, 1968.

18. Bridgman, P. W.
"The Physics of High Pressure"
Bell and Son
London, 1958.

19. Hall, H. T.
Rev. Sci. Inst.
Vol. 31, p. 125.
1960.

20. Blumenthall, G. D.
"Compressive Strength of Hard Metal Alloys"
H.P.M.L. Report, ET67
Johannesburg, 1967.

21. Investigation of Massive Support Principle for Ultra-High Pressure Anvils.
Barogenics Inc.
New York, 1962.

22. Timoshenko, S.
"Strength of Materials, Part II"
Van Nostrand
N.J., 1956.

23. Lengyel, B. and Alexander, J. M.
"Design of a Production Machine for Semi-Continuous Hydrostatic Extrusion"
High Pressure Eng. Conf.
Inst. of Mech. Engrs.
London, 1967.

24. Venter, R. D.
"Optimisation of High Pressure Multi-Component Cylinders"
M.Sc. Thesis, McMaster University
Hamilton, Canada, 1968.

25. Bett, K. E. and Burns, D. J.
"Design of Elastic Multi-Component Compound Cylinders"
High Pressure Eng. Conf.
Inst. Mech. Engrs.
London, Sept. 1967.

26. Becker, S. J. and Mollick, L.
"The Theory of the Ideal Design of a Compound Vessel"
J. Engng. Ind., A.S.M.E.
V. 82 Series B, p. 136.
1960.

27. Parsons, B. and Cole, B. N.
"A Generalised Approach to the Optimum Design of Short Composite Cylinders"
High Pressure Eng. Conf.
Inst. Mech. Engrs.
London, Sept. 1967.

28. Private Communication from
Autoclave Engineers Inc.,
Erie, Pennsylvania.

Harwood Engineering, Co. Inc.
Walpole, Mass.

29. Pippenger, J. J. and Koff, R. M.
"Fluid Power Controls"
McGraw Hill,
N.Y. 1959.

30. Conway, H. G.
"Fluid Pressure Mechanisms"
2nd Ed. Pitmans
London, 1958.

31. Ernst, M. E.
"Oil Hydraulic Power and its Industrial Applications"
McGraw Hill
N.Y. 1949.

32. Leskiewicz, H. J.
"Approach to the Theory of Hydraulic Pressure-Regulating Relief Valves"
A.S.M.E. Paper No. 63
AHGT-47.

33. Merritt, H. E.
"Hydraulic Control Systems"
J. Wiley
N.Y. 1967.

34. Ma, C. Y.
"Analysis and Design of Hydraulic Pressure Reducing Valves"
A.S.M.E. Trans.
J. Engng. Ind. Paper No. 66 - WA/MD-4.
1967.

35. Bialy, L.
"Pressure Control System for a 2000 Ton Experimental Press"
H.P.M.L. Report, D 68.
Johannesburg, 1968.

36. Dowty Technical Information
"Characteristics of Servo-Valves"
Data Sheet 1542
Dowty Rotol Ltd.
Gloucester, 1967.

37. Cooke, C.
"Optimum Pressure for a Hydraulic System"
Product Engineering, p. 162
May 1956.

38. Bialy, L., O'Donovan, K., and de Malherbe, M. C.
"The Continuous Control of High Pressure in Hydraulic Systems"
Proceedings, National Conference on Automatic Control
University of Alberta, Edmonton.
August, 1969.

39. Dorf, R. C.
"Modern Control Systems"
Addison-Wesley
Mass., 1967.

40. Raven, F. H.
"Automatic Control Engineering" (2nd Ed.)
McGraw Hill
N.Y. 1961.

41. Studer, J. J.
"Electronic Circuits and Instrumentation Systems"
J. Wiley
N.Y., 1963.

42. Walters, R.
"Hydraulic and Electro-Hydraulic Servo Systems"
London Iliffe Ltd.
London, 1967.

43. Morrison, J. L. M., Crossland, B., and Parry, J. S. C.
"Strength of Thick Cylinders Subjected to Repeated Internal Pressure"
Proc. Inst. Mech. Engrs. V. 174, p. 95.
London, 1959.

44. Burns, D. J., and Parry, J. S. C.
"Effect of Mean Shear Stress on the Fatigue Behaviour of Thick-Walled Cylinders"
High Pressure Eng. Conf.
Inst. Mech. Engrs.
London, Sept. 1967.

45. Jones, P. M., and Tomkins, B.
"Low Endurance Fatigue of Very Thick-Walled Cylinders"
High Pressure Eng. Conf.
Inst. Mech. Engrs.
London, Sept. 1967.

46. Crossland, B., Skelton, W. J., and Wilson, W. R. D.
"High Pressure Intensifiers:- Development of a Unit for 200,000 p.s.i."
High Pressure Eng. Conf.
Inst. Mech. Engrs.
London, Sept. 1967.

47. Hunsaker, J. C. and Rightmire, B. G.
"Engineering Applications of Fluid Mechanics"
McGraw Hill,
N.Y., 1947.

48. Newhall, D. H. and Abbot, L. H.
"The Bulk Modulus Cell - New High Pressure Measurement Instrument"
Chemical Engineering Progress V. 56, p. 112
Mar. 1960.

49. Dekker, A. J.
"Solid State Physics"
Prentice-Hall
N.J., 1957.

50. Newhall, D. H.
"Manganin High-Pressure Sensors"
Instruments and Control Systems
Nov. 1962.

51. Viersma, T. J.
"Designing Load-Compensated Fast Response Hydraulic Servos"
Control Engng.
May, 1962.

52. Wuest, W.
"Stromung durch Schlitz und Lochblenden bei Kleinen Reynolds-Zahlen"
Ingenieur Archiv. No. 22, p. 357
1954.

53. Timoshenko, S., and Goodier, J. N.
"Theory of Elasticity"
McGraw-Hill
N.Y., 1951.

54. Lewis, E. E., and Stern, H.
"Design of Hydraulic Control Systems"
McGraw-Hill
N.Y., 1962.

55. Dow, R. B.
"The Rheology of Lubricants"
Colloid Sci. V. 2 No. 1
1947.

56. Blackburn, J. F.
"Contributions to Hydraulic Control 5 (Lateral Forces on Hydraulic Pistons)"
A.S.M.E. Trans., V. 75, p. 1175
1953.

57. Sweeney, D. C.
"Preliminary Investigation of Hydraulic Lock"
Engineering, V. 172, p. 513, p. 580.
1951.

58. Manham, J., and Sweeney, D. C.
"An Investigation of Hydraulic Lock"
Proc. Inst. Mech. Engrs.
London, May 13, 1955.

59. Blackburn, J. F., Reethof, G., and Shearer, J. L. (Eds.)
"Fluid Power Control"
Technology Press, M.I.T. and J. Wiley.
N.Y., 1960.

60. Hurwitz, A.
"On the Conditions Under Which an Equation Has Only Roots with Negative Real Parts"
'Selected Papers on Mathematical Trends in Control Theory'
Dover
N.Y., 1964.

61. Cheng, D. K.
"Analysis of Linear Systems"
Addison-Wesley
Mass., 1959.

62. Schultz, W. C. and Rideout, V. C.
"Control System Performance Measures: Past, Present and Future"
I.R.E. Trans. on Automatic Control, p. 22
Feb. 1961.

63. Graham, D. and Lathrop, R. C.
"The Synthesis of Optimum Response: Criteria and Standard Forms, Part 2"
Trans. A.I.E.E., V. 72, p. 273.
Nov. 1953.

64. Schultz, D. G. and Melsa, J. L.
"State Functions and Linear Control Systems"
McGraw Hill
N.Y., 1967.

65. Hohn, F. E.
"Elementary Matrix Algebra" 2nd Ed.
MacMillan
N.Y., 1964.

66. Private Communication from
Industrial Distributors (1946) Ltd.,
Johannesburg
Republic of South Africa.

APPENDIX A

COMPUTER SUBROUTINES AND PROGRAMMES FOR THE TRANSIENT RESPONSE

A-1 SUBROUTINE INIT

A-1-1 Purpose

To set up the coefficients of the characteristic equation of the system (equation (153)); and to establish the matrix, A, in the vector differential equation (165).

A-1-2 Call and Argument List

```
CALL INIT (A, A1, A2, A3, A4, A5, A6, A7, GAIN, G, OM1, OM2,
                                        ZET1, ZET2, H, L)
```

where A(L,L) is the square matrix defined by equation (166) or (167)

L = Total number of phase variables due to the transfer function and the input. (Thus in equation (166), L = 7; in equation (167), L = 8).

A1, A2, ..., A7 are the coefficients a_1, a_2, ..., a_7 of the characteristic equation (153) which are defined by equations (146-1) to (146-7).

GAIN = Forward path gain of the minor loop, denoted k_g in section 6.9-8, and defined by equation (139).

OM1 = Natural frequncy of the servo-valve, denoted ω_1 in section 6.9-7.

OM2 = Natural frequency of the HPA, denoted ω_2 in section 6.9-6, and defined by equation (134).

ZET1 = Damping ratio of the servo-valve, denoted ζ_1 in section 6.9-7.

ZET2 = Damping ratio of the HPA, denoted ζ_2 in section 6.9-6, and defined by equation (135).

H = Displacement transducer feedback gain denoted h in section 6.9-7.

A-1-3 Method

The matrix A(I,J), I = 1,L, J = 1,L is zeroed.
The superdiagonal A(I,I+1), I = 1,L-1 is set equal to unity.
The coefficients A1, A2, ... , A7 are established using equations (146-1) to (146-7). The elements in the sixth row of the matrix are fixed according to equations (166) and (167).

A-1-4 Listing

The listing of SUBROUTINE INIT follows:-

```
      SUBROUTINE INIT(A,A1,A2,A3,A4,A5,A6,A7,GAIN,G,OM1,OM2,ZET1,ZET2,H,
     1L)
C
C     TO ESTABLISH THE COEFFTS OF THE CHARACTERISTIC EQN
C     AND THE VECTOR DIFF EQN MATRIX A(L,L)
C
      DIMENSION A(10,10)
C
C     ZERO THE MATRIX A
C
      DO 200 I=1,L
      DO 200 J=1,L
      A(I,J)=0.
  200 CONTINUE
C
C     ESTABLISH A UNIT SUPER DIAGONAL
C
C     SET UP COEFFTS OF CHAREC. EQN
C
      L1=L-1
      DO 201 I=1,L1
      A(I,I+1)=1.
  201 CONTINUE
C
      A1=1.
      A2=2.*(ZET1*OM1+ZET2*OM2)
      A3=OM1*OM1+OM2*OM2+4.*ZET1*ZET2*OM1*OM2
      A4=2.*OM1*OM2*(ZET1*OM2+ZET2*OM1)
      A5=OM1*OM1*OM2*OM2
      A6=GAIN*A5*H
      A7=G*GAIN*A5
C
C     ESTABLISH REMAINING ELEMENTS OF MATRIX
C
      A(6,1)=-A7
      A(6,2)=-A6
      A(6,3)=-A5
      A(6,4)=-A4
      A(6,5)=-A3
      A(6,6)=-A2
      A(6,7)=+A7
      A(7,8)=1.0
      RETURN
      END
```

A-2 SUBROUTINE CONST

A-2-1 Purpose

To determine whether any absolute stability constraints are violated by the values of the system parameters.

A-2-2 Call and Argument List

CALL CONST (A1, A2, A3, A4, A5, A6, A7, G, H, INDEX)

where INDEX indicates whether or not the system is stable.

The other arguments are defined in A-1-2.

A-2-3 Method

The Routh-Hurwitz criterion (see section 6.12) is used to determine whether any poles of the transfer function lie on the right hand side of the S-plane. The variables B1, B2, C1, C2, E1, F1, denoted as b_1, b_2, c_1, c_2, e_1, f_1 respectively in section 6.12 are defined by equations (154-1) to (154-6).

INDEX = 0 if B1, C1, E1 or F1 is negative, otherwise INDEX = 1.

A unit INDEX indicates stability.

A-2-4 Listing

The listing of SUBROUTINE CONST follows:-

```
      SUBROUTINE CONST(A1,A2,A3,A4,A5,A6,A7,G,H,INDEX)
C
C     TO CHECK FEASIBILITY OF SOLN USING
C     STABILITY AND ACCURACY CRITERIA
C
      INDEX=0
      B1=A3-A1*A4/A2
      B2=A5-A1*A6/A2
      C1=A4-A2*B2/B1
      C2=A6-A2*A7/B1
      E1=B2-B1*C2/C1
      F1=C2-C1*A7/E1
C
C     CONSTRAINTS IMPOSED BY STABILITY REQUIREMENTS
C
      IF(B1.LE.0.)GO TO 2
      IF(C1.LE.0.)GO TO 2
      IF(E1.LE.0.)GO TO 2
      IF(F1.LE.0.)GO TO 2
C
C     CONSTRAINT  IMPOSED BY  ACCURACY REQUIREMENTS
C
      HH=30.*H
      IF(G.LT.HH)GO TO 2
C
C     UNIT INDEX INDICATES FEASIBLE REGION OF SOLUTION
C
      INDEX=1
    2 RETURN
      END
```

A-3 SUBROUTINE OPTRAN

A-3-1 Purpose

(i) To evaluate the transition matrix associated with the solution to the vector differential equation.

(ii) To evaluate the phase variables at preset intervals in time so as to obtain the transient output response of a dynamic system.

(iii) To evaluate the ITAE performance index of the system.

A-3-2 Call and Argument List

CALL OPTRAN (A, B, C, D, X, Y, TINT, TIME, L, NTERM, CRIT)

where A(L,L) is the matrix established by SUBROUTINE INIT

L = Total number of phase variables (see A-1-2)

B(L,L), C(L,L) = Working arrays

D(L,L) = Transition Matrix denoted $\phi(t)$ in section 6.13-2, and defined by equation (169).

Y(L) = Phase variable vector

X(L) = Working vector

TINT = Time interval over which the transiton matrix is to be evaluated, and the intervals at which the phase variables are to be calculated.

TIME = Duration of the transient, an arbitrary finite time chosen so that the ITAE integral, equation (157), approaches a steady-state value. Denoted T_1 in section 6.13-1.

NTERM = The number of time dependent terms in the discrete form of the transition matrix definition, equation (171-1).

Denoted N in section 6.13-3; the series is therefore truncated to NTERM + 1 terms.

CRIT = The value of the ITAE performance index, defined by equation (157).

A-3-3 Method

(i) The transition matrix is established by synthesising equation (171-1). This is effected by evaluating the second term of the series, and then taking the third term to be half the matrix square of the second term, the fourth term is then one third of the matrix product of the second and third terms; the fifth term is one quarter of the product of the second and fourth terms, etc. The terms are then summed and to the sum is added the identity matrix, thus completing the transition matrix. The process is efficiently executed using a series of DO loops, which are documented in the subroutine listing.

(ii) The phase variables are evaluated by applying equation (168). The vector of phase variables after an interval TINT is simply the matrix product of the transition matrix and the initial conditions. The vector thus obtained represents the intial conditions for the next state, i.e. the next interval TINT.

(iii) The ITAE criterion is evaluated using equation (157), which is reproduced below:-

$$\mathrm{ITAE} = \int_{o}^{T_1} t|\varepsilon(t)|\ dt \qquad (157)$$

In FORTRAN notation and in discrete form the absolute error, $\varepsilon(t)$, may be expressed as:-

EP1 = ABS(Y(L) - Y(1))   this is the error at time = TIM

and   EPS = ABS(Y(L) - Y(1))   this is the error at time = TIM - TINT.

The average error from time TIM-TINT to time TIM is thus:-

EP = (EP1 + EPS)/2.0

Equation (157) may thus be expressed as:-

$$\text{CRIT} = \sum_{\text{TIM=TINT}}^{\text{TIME}} (\text{TIM - TINT/2.0})*\text{EP}*\text{TINT}, \text{ where EPS is initially 1.0.}$$

This is easily programmed as shown in the listing.

If an alternative index of performance is to be used, only the FORTRAN statements relevant to the evaluation of the criterion need be altered.

A-3-4 Miscellaneous Comments

(i) Modification for transient response determination

For the purposes of ascertaining the transient response, a WRITE statement may be inserted directly after the statement:- 12 CONTINUE.

The WRITE statement should be of the form:-

```
    WRITE(6, 101)Y(L)
101 FORMAT(1H0, E16.8).
```

(ii) Choosing the value of NTERM, TINT and TIME

It was pointed out in section 6.13-3, that the number of terms and the time interval should be chosen so as to minimise processing time, since subroutine OPTRAN is used many times in the optimisation procedure. Furthermore, the duration of the transient, TIME, should be chosen to be sufficiently long so that CRIT approaches a steady value for each feasible set of parameters, so that a fair discrimination may be made between the values of CRIT calculated for each set of conditions in the optimisation programme.

The best combination of NTERM, TINT and TIME may be rapidly determined using a trial and error method, as follows:-

In section 6.13-3, it was shown that the time interval TINT (denoted T in 6.13-3) should be in the range $\frac{1}{a_2} \leq \mathrm{TINT} \leq \frac{7}{a_2}$.

And since a_2 is of the order 20,000; $5 \times 10^{-5} \leq \mathrm{TINT} \leq 3.5 \times 10^{-4}$ secs.

The factor $\frac{1}{\mathrm{NTERM!}}$ (denoted $\frac{1}{N!}$ in equation (171-1)) decreases very rapidly as NTERM is increased.

For NTERM = 10, $\frac{1}{\mathrm{NTERM!}} \simeq 0.28 \times 10^{-6}$;

and for NTERM = 15, $\frac{1}{\mathrm{NTERM!}} \simeq 0.76 \times 10^{-12}$;

while for NTERM = 20, $\frac{1}{\mathrm{NTERM!}} \simeq 0.41 \times 10^{-18}$.

If NTERM is chosen as 15 the value of TINT may be decreased steadily from its maximum value until the transition matrix converges. This entails the inclusion of a WRITE statement immediately after statement 8 in the subroutine, to print the transition matrix. For the arbitrary values of G = 200, H = 1, a value of TINT $\equiv 7.5 \times 10^{-5}$ was found to be sufficiently small to cause all the elements in the transition matrix to converge to 8 decimal places.

Using the same arbitrary values of G and H, and with TINT = 7.5×10^{-5}, NTERM = 15, CRIT is determined from OPTRAN for successively increasing values of TIME until CRIT has converged to 8 decimal places. This occurs when TIME = 0.1 sec. Using TIME = 0.1 sec, the values of NTERM and TINT are adjusted successively so as to achieve the same convergence of the transiton matrix as previously, until the values which minimise the processing time are found. This was found to occur when NTERM = 17, TINT = 1×10^{-4} sec.

A-3-5 Listing

The listing of SUBROUTINE OPTRAN is as follows:-

```
      SUBROUTINE OPTRAN(A,B,C,D,X,Y,TINT,TIME,L,NTERM,CRIT)
C
C     TO DETERMINE
C                    (1) THE TRANSITION MATRIX D(L,L)
C                    (2) THE ITAE CRITERION, CRIT
C                    (3) THE PHASE VARIABLES Y(L)
C
      DIMENSION A(10,10),B(10,10),X(10),Y(10),C(10,10),D(10,10)
C
C     ESTABLISH 2ND TERM OF EXPONENTIAL SERIES
C
      DO 1 I=1,L
      DO 1 J=1,L
      A(I,J)=A(I,J)*TINT
    1 CONTINUE
C
C     SETUP WORKING MATRICES B,C. AND PUT D=2ND TERM.
C
      DO 2 I=1,L
      DO 2 J=1,L
      B(I,J)=A(I,J)
      D(I,J)=A(I,J)
      C(I,J)=0.
    2 CONTINUE
C
C     ESTABLISH OTHER TERMS IN THE SERIES
C
      DO 3 NDEM=2,NTERM
      DEN=NDEM
      DO 4 I=1,L
      DO 4 J=1,L
      DO 4 K=1,L
C
C     ESTABLISH ATERM OF THE SERIES
C
      C(I,J)=C(I,J)+A(I,K)*B(K,J)/DEN
    4 CONTINUE
      DO 5 I=1,L
      DO 5 J=1,L
C
C     RETAIN PREVIOUS TERM,C,AS B TO FORM BASIS OF NEXT TERM
C
      B(I,J)=C(I,J)
C
C     SUM THE SERIES ,LET D =SUM
C
      D(I,J)=D(I,J)+C(I,J)
    5 CONTINUE
```

```
C
C     WORKING MATRIX C IS NOW ZEROED IN PREPARATION FOR
C     ANOTHER PASS THROUGH THE DO LOOP
C
      DO 6 I=1,L
      DO 6 J=1,L
      C(I,J)=0.
    6 CONTINUE
    3 CONTINUE
C
C     SERIES IS NOW COMPLETE EXCEPT FOR THE FIRST TERM
C     C(I,J) IS REDEFINED AS THE IDENTITY MATRIX
C
      DO 7 I=1,L
      C(I,I)=1.
    7 CONTINUE
      DO 8 I=1,L
      DO 8 J=1,L
C
C     THE SUMMATION IS CONCLUDED
C
      D(I,J)=D(I,J)+C(I,J)
    8 CONTINUE
C
C     D(I,J) IS THE TRANSITION MATRIX
C
C     CALCULATE STATE VARIABLES
C
      TIM=TINT
      CRIT=0.
      EPS=1.
    9 CONTINUE
      DO 10 I=1,L
      X(I)=0.
   10 CONTINUE
      DO 11 I=1,L
      DO 11 J=1,L
      X(I)=X(I)+D(I,J)*Y(J)
   11 CONTINUE
C
C     RETAIN CURRENT OUTPUT AS INITIAL CONDITION FOR THE NEXT STATE
C
      DO 12 I=1,L
      Y(I)=X(I)
   12 CONTINUE
      EP1=ABS(Y(L)-Y(1))
      EP=(EP1+EPS)/2.
      DEL=(TIM-TINT/2.)*EP*TINT
      CRIT=CRIT+DEL
      EPS=EP
      IF(TIM.GE.TIME)RETURN
      TIM=TIM+TINT
      GO TO 9
      END
```

A-4 PROGRAMME TRANSIM

A-4-1 Purpose

To simulate the response of the system to a ramp-plateau negative ramp cycle.

A-4-2 Method

The intial conditions are read in as YY(I), I = 1,L. These represent the initial conditions of the state vector Y(L). The three stages of the simulation are conducted in sequence, NSIM being used as a counter to determine the appropriate value L, of the vector length and to modify the initial conditions of Y(7) and Y(8) at the commencement of each new segment of the cycle. The other initial conditions are merely carried over from the preceding segment (i.e. the final conditions of the first segment become the initial conditions of the second segment, etc.).

A-4-3 Subroutines Used

INIT is used to set up the matrix A, and evaluate A1, A2, ..., A7 for subroutine CONST.

CONST confirms the stability of the system and hence the feasibility of the parameters.

OPTRAN calculates the transient response.

A-4-4 Miscellaneous Comments

To simulate the response for example to a step input only, a similar programme would be used, with the first and the third stages of the simulation omitted.

A-4-5 Listing

The listing of PROGRAMME TRANSIM follows:-

```
C       PROGRAMME TRANSIM
C
C       FOR DETERMINING TRANSIENT RESPONSE USING PHASE VARIABLES
C
C       SIMULATION OF RESPONSE TO A RAMP-PLATEAU-NEGATIVE RAMP CYCLE
C
        DIMENSION A(10,10),B(10,10),X(10),Y(10),C(10,10),D(10,10),YY(10)
        READ(5,101)TINT,L,NTERM,TIME
        READ(5,102)(YY(I),I=1,L)
        READ(5,110)GAIN,G,OM1,OM2,ZET1,ZET2,H
        WRITE(6,620)
        WRITE(6,621)TINT,L,NTERM,TIME
        WRITE(6,622)
        WRITE(6,623)(YY(I),I=1,8)
        WRITE(6,624)
        WRITE(6,625)GAIN,G,OM1,OM2,ZET1,ZET2,H
C
C         SET  Y  TO THE SYSTEM INPUT YY
C
        DO 300 I=1,L
        Y(I)=YY(I)
   300  CONTINUE
        WRITE(6,103)
C
C       FIRST STAGE OF SIMULATION 100 PSI/SEC RAMP FOR 100MS.
C
        NSIM=0
   302  CONTINUE
        CALL INIT(A,A1,A2,A3,A4,A5,A6,A7,GAIN,G,OM1,OM2,ZET1,ZET2,H,L)
C
C       TEST THAT INITIAL POINT IN FEASIBLE REGION
C
        CALL CONST(A1,A2,A3,A4,A5,A6,A7,G,H,INDEX)
        IF(INDEX.EQ.1)GO TO 301
        WRITE(6,100)
        STOP
   301  CONTINUE
        CALL OPTRAN(A,B,C,D,X,Y,TINT,TIME,L,NTERM,CRIT)
        NSIM=NSIM+1
        GO TO(303,304T305)NSIM
C
C       2ND. STAGE OF SIMULATION 10 PSI PLATEAU FOR 100 MS.
C
   303  Y(7)=10.
        Y(8)=0.
        L=7
        GO TO 302
```

```
C
C      3RD. STAGE OF SIMULATION -100 PSI/SEC RAMP FOR 100 MS.
C
  304 CONTINUE
      Y(8)=-100.
      L=8
      GO TO 302
  305 STOP
 100  FORMAT(39H INITIAL POINT OUTSIDE FEASIBLE REGION ///)
  101 FORMAT(E10.3,2I3,E10.3)
  102 FORMAT(8F6.1)
  103 FORMAT(1H0,*   TIME (SEC)         INPUT               OUTPUT*,//)
  110 FORMAT(7F10.3)
  620 FORMAT(1H0,*TINT,L,NTERM,TIME*)
  621 FORMAT(1H0,E10.3,2I3,E10.3)
  622 FORMAT(1H0,*INPUT, YY=*)
  623 FORMAT(1H0,F6.1)
  624 FORMAT(1H0,*GAIN,G,OM1,OM2,ZET1,ZET2,H*)
  625 FORMAT(1H0,7F10.3)
      STOP
      END
```

APPENDIX B

DIRECT-SEARCH OPTIMISATION PROGRAMME FOR THE CONTROL SYSTEM

B-1 Purpose

To find the values of the parameters G and H (denoted K and h in Chapter 6) which optimise the transient performance of the control system.

B-2 Notation

VO = 'Normalised' variable proportional to G i.e. VO = G/300.0

WO = 'Normalised' variable proportional to H i.e. WO = H/1.4

R = Length of the search vector in VO - WO space.

THETA = Angle the search vector makes with the VO axis.

P = Rotation index

Q = Rotational incrementing index

V = Temporary value of VO

W = Temporary value of WO

U = Temporary value of CRIT.

The other variables used have been defined in Appendix A.

B-3 Method

The direct search method used is a 'polar' technique of searching rather than the more usual 'cartesian' type. Thus the search vector may be rotated through a full circle by variable increments, and may be increased or decreased in length as required. The search vector changes its origin as the optimisation procedes, as in the case of the conventional direct search.

The advantage of this technique is that it prevents the search from stalling on a 'fence' at sub-optimum values of the optimisation function.

Starting from an initial feasible point which is confirmed by Subroutine CONST, (see flow chart, Fig. 50), the value of CRIT is calculated by Subroutine OPTRAN and temporarily stored as U. A step of length R is taken in a direction THETA to the VO axis (i.e. the G axis).

The angle THETA is varied by the parameters P and Q, according to the relationship:-

THETA = 0.7853982*(1.0 + P/Q)

Thus when P = 0.0, THETA would be 45°. When P = 8.0*Q, a full rotation has occurred. Q is used to alter the increment in THETA. If for example Q = 12.0, the increments in THETA as P is varied by units 1.0 would be 3.75°.

The new position of the search point is stored as V,W. The feasibility of this point is again assessed using subroutines INIT and CONST and if feasible, OPTRAN is again called upon to calculate the value of CRIT. If an improvement has occurred over the preceding value, the step length is increased and the procedure repeated. If the new point is infeasible or no improvement occurs, the step length, if larger than its original starting value is reset to the starting value, and P is increased causing the search vector to rotate. The procedure is then repeated.

When a full rotation of the search vector has taken place, the increment on rotation is reduced, and the step length decreased until the optimum point is reached to within an acceptable tolerance.

For the initial optimisation, the initial value of R used was 0.1, and the programme was stopped when R $\leqslant$ 0.01, and Q = 12.0 (a limit statement prevents Q from exceeding 12.0). For the final optimisation the initial value of R was 0.01, and the stopping value 0.0005, while Q was limited to 45.0. (See section 6.13-4.)

B-4 Miscellaneous Comments

The 'polar' direct search technique may be extended to three or more variables using spherical or n-dimensional polar space as required.

B-5 Listing

The listing of the initial optimisation programme follows:-

```
C     OPTIMISATION OF TRANSIENT RESPONSE USING PHASE VARIABLES AND A
C     DIRECT SEARCH TECHNIQUE
C
      DIMENSION A(10,10),B(10,10),X(10),Y(10),C(10,10),D(10,10),YY(10)
      READ(5,101)TINT,L,NTERM,TIME
      READ(5,102)(YY(I),I=1,L)
      READ(5,110)GAIN,G,OM1,OM2,ZET1,ZET2,H
      WRITE(6,620)
      WRITE(6,621)TINT,L,NTERM,TIME
      WRITE(6,622)
      WRITE(6,623)(YY(I),I=1,8)
      WRITE(6,624)
      WRITE(6,625)GAIN,G,OM1,OM2,ZET1,ZET2,H
      Q=1.
      P=0.
      W0=H/1.4
      V0=G/300.
      R=0.1
      CALL INIT(A,A1,A2,A3,A4,A5,A6,A7,GAIN,G,OM1,OM2,ZET1,ZET2,H,L)
C
C     TEST THAT INITIAL POINT IN FEASIBLE REGION
C
      CALL CONST(A1,A2,A3,A4,A5,A6,A7,G,H,INDEX)
      IF(INDEX.EQ.1)GO TO 301
      WRITE(6,100)
      STOP
  301 CONTINUE
C
C       SET  Y  TO THE SYSTEM INPUT YY
C
      DO 300 I=1,L
      Y(I)=YY(I)
  300 CONTINUE
      CALL OPTRAN(A,B,C,D,X,Y,TINT,TIME,L,NTERM,CRIT)
      WRITE(6,502)
      WRITE(6,501)G,H,CRIT
  302 U=CRIT
  303 THETA=0.7853982*(1.+P/Q)
C
C     I.E.  THETA = 45+45*P/Q
C
C     TRANSFORM FROM POLAR TO CARTESIAN COORDS.
C
      V=V0+R*COS(THETA)
      W=W0+R*SIN(THETA)
      G=300.*V
      H=1.4*W
```

```
C
C     SEARCH FOR ANEW BASE POINT.
C
      CALL INIT(A,A1,A2,A3,A4,A5,A6,A7,GAIN,G,OM1,OM2,ZET1,ZET2,H,L)
      CALL CONST(A1,A2,A3,A4,A5,A6,A7,G,H,INDEX)
C
C     EXAMINE WHETHER AN IMPROVEMENT HAS OCCURED
C
      IF(INDEX.NE.1)GO TO 304
C
C     RESET  Y  TO THE SYSTEM INPUT YY
C
      DO 307 I=1,L
      Y(I)=YY(I)
  307 CONTINUE
      CALL OPTRAN(A,B,C,D,X,Y,TINT,TIME,L,NTERM,CRIT)
      IF(CRIT.GE.0)GO TO 304
C
C     IF IMPROVEMENT , VO, WO, NEW BASE POINT
C
      VO=V
      WO=W
      WRITE(6,501)G,H,CRIT
C
C     THE RADIUS VECTOR IS INCREASED WHILE THE GOING IS GOOD
C     (WITHOUT ALTERING THETA)
C
      R=1.5*R
      GO TO 302
  304 QQ=8.*Q
C
C     IF FAILURE OCCURS AFTER A RUN OF SUCCESSES,
C     R IS RESET TO ITS ORIGINAL VALUE OF 0.1  (ONLY APPLIED IF R.GT.0.1)
C
      IF(R.GT.0.1)R=0.1
      IF(P.GE.QQ)GO TO 305
C
C     THE RADIUS VECTOR IS ROTATED BY  INCREMENTING P WHICH ALTERS THETA
C
      P=P+1.
      GO TO 303
C
C     IF NO IMPROVMENT AFTER COMPLETE REVOLUTION OF RADIUS,
C      R AND THETA ARE REDUCED (I.E. INCS. IN THETA ARE REDUCED)
C     THE PROCESS IS CONTINUED UNTIL INCS. IN THETA ARE 3.75 DEGREES
C      THERAFTER R IS REDUCED  WITH EACH CYCLE UNTIL R.LE.1.E-2
C      THIS TERMINATES THE PROGRAMME.
C
  305 IF(R.LE.1.E-2)GO TO 306
```

```
      R=0.5*R
      Q=Q+1.
      IF(Q.GE.13.)Q=12.
      P=0.
      GO TO 303
  306 WRITE(6,501)G,H,U
      WRITE(6,626)
  100 FORMAT(39H INITIAL POINT OUTSIDE FEASIBLE REGION ////)
  101 FORMAT(E10.3,2I3,E10.3)
  102 FORMAT(8F6.1)
  110 FORMAT(7F10.3)
  501 FORMAT(1H0,2F16.10,E16.8)
  502 FORMAT(1H0,*             G                    H                 CRIT*,///)
  620 FORMAT(1H0,*TINT,L,NTERM,TIME*)
  621 FORMAT(1H0,E10.3,2I3,E10.3)
  622 FORMAT(1H0,*INPUT, YY=*)
  623 FORMAT(1H0,F6.1)
  624 FORMAT(1H0,*GAIN,G,OM1,OM2,ZET1,ZET2,H*)
  625 FORMAT(1H0,7F10.3)
  626 FORMAT(1H0,*OPTIMUM HAS BEEN REACHED *)
      STOP
      END
```

APPENDIX C

COMPUTER PROGRAMME AND SUBROUTINES FOR SIMULATING THE FREQUENCY RESPONSE

C-1 PROGRAMME FREQRES

C-1-1 Purpose

To determine the frequency response of most standard forms of open loop transfer functions.

C-1-2 Notation

The terms used as variables in FREQRES and as arguments in the associated subroutines are defined in the listing of programme FREQRES.

C-1-3 Method

Programme FREQRES is intended to cope with transfer functions of the general form:- (written in algebraic notation)

$$\frac{\mathrm{GAIN}.S^{N}(1 + S/\mathrm{FREN})(1 + 2(\mathrm{ZEN})S/\mathrm{OMEN} + S^2/(\mathrm{OMEN})^2)(1 + C(1)S + C(2)S^2 + \ldots + C(\mathrm{NN})S^{\mathrm{NN}}}{(1 + S/\mathrm{FRED})(1 + 2(\mathrm{ZED})S/\mathrm{OMED} + S^2/(\mathrm{OMED})^2)(1 + \mathrm{CD}(1)S + \mathrm{CD}(2)S^2 + \ldots + \mathrm{CD}(\mathrm{ND})S^{\mathrm{ND}})}$$

where (1 + S/FREN) represents all the linear terms in the numerator,

i.e. $(1 + S/\mathrm{FREN}) = (1 + S/\mathrm{FREN}(1))(1 + S/\mathrm{FREN}(2)) \ldots (1 + S/\mathrm{FREN}(L))$

also $(1 + S/\mathrm{FRED}) = (1 + S/\mathrm{FRED}(1))(1 + S/\mathrm{FRED}(2)) \ldots (1 + S/\mathrm{FRED}(M))$

$$(1 + 2S(\mathrm{ZEN})/\mathrm{OMEN} + S^2/(\mathrm{OMEN})^2) = (1 + 2S(\mathrm{ZEN}(1))/\mathrm{OMEN}(1) + S^2/(\mathrm{OMEN}(1))^2)\ldots$$
$$\ldots (1 + 2S(\mathrm{ZEN}(\mathrm{LL}))/\mathrm{OMEN}(\mathrm{LL}) + S^2/(\mathrm{OMEN}(\mathrm{LL}))^2)$$

And a similar expression for the quadratic term in the denominator, where OMED has the values OMED(1) ... OMED(MM), and ZED has the values ZED(1) ... ZED(MM).

Terms of the form S^N are processed using Subroutine ESS. Linear, quadratic and polynomial terms are processed by subroutines LIN, QUAD and POLY respectively.

The programme uses the values of the variables N, L, M, LL, MM, NN, ND to determine which subroutines are to be used and how often each is to be applied. The magnitudes are multiplied together and the phase angles added after each subroutine returns to the main programme.

C-1-4 Users Information

The variables N, L, M, LL, MM, NN, ND, LIM are punched on the first card in 8I4 Format. (In the absence of one or more of these terms a zero is punched.)

The other cards (in order) contain the following data, where relevant:-

FREN(I)	Format F 16.8, one value per card.
FRED(I)	Format F 16.8, one value per card.
ZEN(I), OMEN(I)	Format 2F 16.8 one card for each value of I.
ZED(I), OMED(I)	Format 2F 16.8 one card for each value of I.
C(I)	Format F 16.8 one value per card.
CD(I)	Format F 16.8 one value per card.
OMEGA	Format F 16.8 one value per card.

Note:- The number of frequncy values, OMEGA, to be processed is LIM. The frequencies should be in units of Hertz.

C-2 THE SUBROUTINES

The output arguments of Subroutines ESS, LIN, QUAD and POLY are the magnitude, A, and phase angle FI. The input arguments appear in the CALL list of each subroutine and are defined in the listing of programme FREQRES.

Subroutines ESS and LIN are straightforward and their operation is clear from the listing. Subroutines QUAD and POLY warrant a brief discussion.

C-2-1 SUBROUTINE QUAD

Two special cases arise in a second order transfer function. Firstly, if the damping ratio is zero, the magnitude is given by:-

A = ABS(1.0 - R*R) where R = OMEGA/OM = frequency ratio.

The phase angle is zero if 1.0 > R*R, and 180° if R*R > 1.0.

Secondly, at the natural frequency, the second order transfer function becomes a pure integrator (if in the denominator) or differentiator (if in the numerator), which is amplified by the term 2.0*ZETA. This is easily dealt with as in ESS.

The general case is easily handled by recognising that the transfer function may be written as:- X + JY where J is the complex operator, X = 1.0 - R*R, Y = 2.0*ZETA*R.

Thus A = SQRT(X*X + Y*Y), and the phase angle is determined from the arc tangent of the complex expression.

C-2-2 SUBROUTINE POLY

The polynomial in the laplace operator S is partitioned into four groups, P1, P2, P3, P4 corresponding to phase angles of 0°, 90°, 180° and 270° respectively. The grouping is efficiently carried out by making N/3 passes through a DO loop. (If N/3 is less than 2, two passes are made as this is the minimum number of passes required to complete the partitioning (see listing).)

By putting X = P1 - P3, and Y = P2 - P4, a form identical to that obtained in the general case of subroutine QUAD results and is processed in the same manner.

C-3 THE LISTING

The listing of programme FREQRES and the various subroutines follows:-

```
C       PROGRAMME FREQRES
C
C       PROGRAMME TO EVALUATE OPEN -LOOP TRANSFER FUNCTIONS
C
C       GAIN = O/L GAIN OF TRANSFER FUNCTION
C       N=POWER OF S  (-1 FOR 1 INTEGRATION, +2 FOR 2 DIFFERATIONS ETC.)
C       L=NO. OF LINEAR TERMS  (1.+S/FREN) IN NUMERATOR
C       M=NO. OF LINEAR TERMS  (1.+S/FRED) IN DENOMENATOR
C        FREN, FRED =NAT. FREQ. OF LINEAR TERMS
C       OMEN,OMED=NAT FREQ. OF QUAD. TERMS
C       LL= NO. OF QUAD. TERMS  (1.+2.*ZEN*S/OMEN+S**2/OMEN**2) IN NUM.
C       MM= NO. OF QUAD. TERMS  (1.+2.*ZED*S/OMED+S**2/OMED**2) IN DEN.
C       ZEN,ZED=DAMPING RATIOS IN NUM AND DEN RESPLY.
C       NN=ORDER OF POLY IN S IN NUM I.E. 1+C(1)*S+C(2)*S**2+...+C(NN)*S**NN
C       NN=ORDER OF POLY IN S IN DEN I.E. 1+CD(1)*S+...+CD(ND)*S**ND
C       ND IS ORDER OF POLY IN DENOM.
C       LIM=NO. OF FORCING FREQ. POINTS TO BE EXAMINED
C       OMEGA=FORCING FREQ.
C       OMEGA IS READ IN AS HERTZ
C
C
        DIMENSIONFREN(20),ZEN(20),OMEN(20),FRED(20),ZED(20),OMED(20),C(50)
       1,CD(50)
C
C       DATA DESCRIBING FORM OF TRANS. FN. IS READ IN
C
        READ(5,101)N,L,M,LL,MM,NN,ND,LIM
        WRITE(6,556)N,L,M,LL,MM,NN,ND,LIM
C
C       DETERMINE IF ANY LINEAR TERMS ARE PRESENT
C
        IF(L.EQ.0)GO TO 11
        READ(5,102)(FREN(I),I=1,L)
     11 IF(M.EQ.0)GO TO 12
        READ(5,102)(FRED(I),I=1,M)
C
C       DETERMINE IF ANY  QUAD. TERMS ARE PRESENT
C
     12 IF(LL.EQ.0)GO TO 13
        READ(5,103)(ZEN(I),OMEN(I),I=1,LL)
     13 IF(MM.EQ.0)GO TO 14
        READ(5,103)(ZED(I),OMED(I),I=1,MM)
C
C       DETERMINE IF ANY  POLY  TERMS ARE PRESENT
C
     14 IF(NN.EQ.0)GO TO 15
        READ(5,102)(C(I),I=1,NN)
     15 IF(ND.EQ.0)GO TO 19
        READ(5,102)(CD(I),I=1,ND)
     19 CONTINUE
```

```
      WRITE(6,120)
      KOUNT=0
   17 READ(5,102)OMEGA
C
C     OMEGA IS CONVERTED TO RADS./SEC
C
      OMEGA=OMEGA*6.2831853
   18 CONTINUE
      KOUNT=KOUNT+1
      AMAG=GAIN
      PHASE=0.
      IF(N.EQ.0)GO TO 20
C
C     IF ANY INTEGRATIONS OR DIFFERENTIATIONS ARE REQD. THEY ARE PERFORMED
C     USING SUBROUTINE ESS
C
      CALL ESS(N,OMEGA,A,FI)
      AMAG=AMAG*A
      PHASE=PHASE+FI
   20 CONTINUE
      IF(L.EQ.0)GO TO 21
      DO 21 J=1,L
      FREQ=FREN(J)
C
C     LINEAR TERMS ARE SOLVED
C
      CALL LIN(OMEGA,FREQ,A,FI)
      AMAG=AMAG*A
      PHASE=PHASE+FI
   21 CONTINUE
      IF(M.EQ.0)GO TO 22
      DO 22 J=1,M
      FREQ=FRED(J)
      CALL LIN(OMEGA,FREQ,A,FI)
      AMAG=AMAG/A
      PHASE=PHASE-FI
   22 CONTINUE
      IF(LL.EQ.0)GO TO 23
      DO 23 J=1,LL
      ZETA=ZEN(J)
      OM=OMEN(J)
C
C     QUAD.  TERMS ARE SOLVED
C
      CALL QUAD(OMEGA,OM,ZETA,A,FI)
      AMAG=AMAG*A
      PHASE=PHASE+FI
   23 CONTINUE
```

```
      IF(MM.EQ.0)GO TO 24
      DO 24 J=1,MM
      ZETA=ZED(J)
      OM=OMED(J)
      CALL QUAD(OMEGA,OM,ZETA,A,FI)
      AMAG=AMAG/A
      PHASE=PHASE-FI
   24 CONTINUE
      IF(NN.EQ.0)GO TO 25
C
C     POLY.  TERMS ARE SOLVED
C
      CALL POLY(OMEGA,C,NN,A,FI)
      AMAG=AMAG*A
      PHASE=PHASE+FI
   25 CONTINUE
      IF(ND.EQ.0)GO TO 26
      CALL POLY(OMEGA,CD,ND,A,FI)
      AMAG=AMAG/A
      PHASE=PHASE-FI
   26 CONTINUE
C
C     AMAG=MAGNITUDE ,PHASE=PHASE LEAD   (IF -VE THEN LAG)
C     DB=POWER GAIN (DECIBELS)
C
C    PHASE. IS CONVERTED TO DEGREES
C
      PHASE=180.*PHASE/3.14159265
C
C     MAG.. IS CONVERTED TO DECIBELS
C
      DB=20.*ALOG10(AMAG)
C
C     FREQ. IS CONVERTED TO HERTZ
C
      OMEGA=OMEGA/6.2831853
      WRITE(6,121)OMEGA,AMAG,DB,PHASE
C
C     PROGRAMME IS TERMINATED WHEN ALL FORCING FREQS. HAVE BEEN PROCESSED
C
      IF(KOUNT.GE.LIM)STOP
      GO TO 16
  101 FORMAT(8I4)
  102 FORMAT(F16.8)
  103 FORMAT(2F16.8)
  120 FORMAT(1H0,* FREQ.( HERTZ )     MAGNITUDE     MAG.(DECIBELS)  PHASE
     1(DEGREES)*,//)
  121 FORMAT(1H0,4F16.8)
  556 FORMAT(1H0,*DATA*,8I4,//)
      END
```

```
      SUBROUTINE ESS(N,OMEGA,A,FI)
C
C     TO SOLVE LAPLACE FUNCTION OF THE FORM S**N
C
      NA=-N
      IF(N.GT.0)GO TO 1
      A=1./OMEGA**NA
      GO TO 2
    1 CONTINUE
      A=OMEGA**N
    2 CONTINUE
      EN=N
      FI=EN*1.57079633
      RETURN
      END
```

```
      SUBROUTINE LIN(OMEGA,FREQ,A,FI)
C
C     TO SOLVE LAPLACE FUNCTIONS OF THE FORM
C     (1.+S/FREQ) AND (1.-S/FREQ)
C
      R=OMEGA/FREQ
C
C     CALCULATE MAGNITUDE
C
      A=SQRT(1.+R*R)
      IF(R.LT.0.)GO TO 1
C
C     DETERMINE PHASE ANGLE FOR (1.+S/FREQ) CASE
C
      FI=ATAN(R)
      RETURN
C
C     DETERMINE PHASE ANGLE FOR (1.-S/FREQ) CASE
C
    1 FI=6.28318531-ATAN(ABS(R))
      RETURN
      END
```

```
      SUBROUTINE QUAD(OMEGA,OM,ZETA,A,FI)
C
C     FOR LAPLACE FUNCTIONS OF THE GENERAL FORM
C     (1.+2.*ZETA*S/OM+S*S/(OM*OM))
C     OM=NATURAL FREQ.,  ZETA=DAMPING RATIO,  OMEGA=FORCING FREQ.
C
      R=OMEGA/OM
C
C     SPECIAL CASE 1 - ZERO DAMPING
C
      IF(ZETA.NE.0.)GO TO 1
      X=1.-R*R
      IF(X.GE.0.)FI=0.
      IF(X.LE.0.)FI=3.14159265
      A=ABS(X)
      RETURN
C
C     SPECIAL CASE 2 - AT NATURAL FREQUENCY
C
    1 IF(OMEGA.NE.OM)GO TO 2
      Y=2.*ZETA*R
      IF(Y.GT.0.)FI=1.57079633
      IF(Y.LT.0.)FI=4.71238898
      A=ABS(Y)
      RETURN
C
C     THE GENERAL CASE
C
    2 X=1.-R*R
      Y=2.*ZETA*R
C
C     DETERMINE MAGNITUDE AND PHASE
C
      A=SQRT(X*X+Y*Y)
      ANG=ATAN(ABS(Y/X))
C
C     LOCATE QUADRANT
C
      IF(X.GT.0.0.AND.Y.GT.0.)FI=ANG
      IF(X.GT.0.0.AND.Y.LT.0.)FI=6.28318531-ANG
      IF(X.LT.0.0.AND.Y.GT.0.)FI=3.14159265-ANG
      IF(X.LT.0.0.AND.Y.LT.0.)FI=3.14159265+ANG
C
      RETURN
      END
```

```
      SUBROUTINE POLY(OMEGA,C,N,A,FI)
C
C     TO SOLVE POLYS IN THE LAPLACE OPERATOR , S
C     POLY OF THE FORM  1.0+C(1)*S+C(2)*S*S+C(3)*S**3+....+C(N)*S**N
C
C     THE TERMS ARE GROUPED ALONG THE 4 VECTORIAL DIRECTIONS
C     OF THE ARGAND DIAGRAM
C
C     FOR N.LT.6 AT LEAST TWO PASSES THROUGH A DO LOOP ARE REQD.
C     OTHERWISE AT LEAST N/3 PASSES ARE REQD.
C
C     P1,P2,P3,P4 ARE SUMS OF PARTITIONED POLYS CORRESP. TO EACH DIRECTION
C
      DIMENSIONC(50)
      P1=1.
      P2=0.
      P3=0.
      P4=0.
      M=N/3
      IF(M.LE.2)M=2
      DO 1 I=1,M
C
C     GROUPING
C
      IF((N-(4*I)).GE.0)P1=P1+C(4*I)*OMEGA**(4*I)
      IF((N-(4*I)+3).GE.0)P2=P2+C(4*I-3)*OMEGA**(4*I-3)
      IF((N-(4*I)+2).GE.0)P3=P3+C(4*I-2)*OMEGA**(4*I-2)
      IF((N-(4*I)+1).GE.0)P4=P4+C(4*I-1)*OMEGA**(4*I-1)
    1 CONTINUE
C
C     PUT IN COMPLEX FORM,  I.E. X+JY
C
      X=P1-P3
      Y=P2-P4
C
C     DETERMINE MAGNITUDE AND PHASE
C
      A=SQRT(X*X+Y*Y)
      ANG=ATAN(ABS(Y/X))
C
C     LOCATE QUADRANT
C
      IF(X.GT.0.0.AND.Y.GT.0.)FI=ANG
      IF(X.GT.0.0.AND.Y.LT.0.)FI=6.28318531-ANG
      IF(X.LT.0.0.AND.Y.GT.0.)FI=3.14159265-ANG
      IF(X.LT.0.0.AND.Y.LT.0.)FI=3.14159265+ANG
C
      RETURN
      END
```