NIKHEF

NATIONAAL INSTITUUT VOOR KERNFYSICA EN HOGE-ENERGIEFYSICA

NIKHEF - K - MTG - 21

A.P. KAAN, J.F.J. VAN DEN BRAND

DESIGN STUDY OF THE VACUUM SYSTEM OF THE E.S.R.P.

AFDELINGSRAPPORT
NIKHEF - K -
MTG-21

Abstract

The vacuum system of ESRP is studied by means of a static and dynamic finite element calculation method. The remarkable result appeared to be the possibility of pumping this system only by lumped pumps, by localizing most of the radiation induced gas load at a number of special absorbers ("crotches").

1. Introduction

In order to contribute to the design of the ESRP we made a study of the vacuum system. This paper is an update and a more comprehensive version of the April-report.
The results are based on the construction information we got from J. Bijleveld*. Further investigations are necessary when "the geometry of the vacuum envelope" is final in order to obtain precise average pressures. Special attention must be given to the pressure profile in the wigglers, undulators, cavities, septum magnet, etc.

The most important design criterium for the vacuum system is the beam lifetime. We assumed a beam lifetime of > 10 h is desired. Gasscattering is varying approximately with Z^{+2}. The required partial pressures are listed in ESRP supplement II: "The Machine". In practice the ratios of the partial pressures are H2/CO = 3-4, CO_2/CO = 0.2 - 0.7 (stored beam conditions) and for this reason our calculations are based on a total pressure of 1.10^{-9} Torr for mass 28.
We assumed the all metal vacuum envelope was constructed of stainless steel (no organic materials) with conflat-type sealings. Before (clean room) assembly all the vacuum components have to be chemically cleaned (no acids) and vacuum baked (to 900°C) resulting in a desorption rate of about 1.10^{-12} Torr $ls^{-1}cm^{-2}$. Bake-out in situ up to 300°C has to be possible in order to achieve low desorption rates (no argon in situ glow discharge).

* ESRP IRM 47/84

CONTENTS

The calculation was based on a desorption rate of 3×10^{-12} Torr $ls^{-1}cm^{-2}$ in this way taking a safety factor of 3 into account. During operation of the ESRP the predominant gas load is caused by sunchrotron radiation induced gas desorption. An estimate of this gas load can be made on basis of the expected photo electron current and the number of desorbed molecules per electron. A calculation (ESRP-IRM-84 S, Tazzari) gives a total gas load for the ESRP due to synchrotron radiation of approximately 1×10^{-5} Torr ls^{-1} (E = 6 GeV, I = 0.2 A after 150 AH beam cleaning. In our calculations we used a gas load of 2×10^{-7} Torr ls^{-1} for each bending magnet distributed over three crotches.
It is assumed that 10% of the radiation is uniformly distributed in the magnet chambers.

This paper is organized as follows: the mathematical theory is presented in the appendix and is applied to vacuum systems in chapter 2.
The ESRP-case is studied in chapter 3. Finally the cost estimates are given in chapter 4.

2. Calculation of the characteristics of vacuum systems using finite-element analysis

It will be discussed how a pressure-distribution in a vacuum system can be calculated. The pressure distribution depends on the geometry of the system, the desorption, the pumping speed and the time. The vacuum system is divided in a number of elements. In each element the pressure is calculated taken into account the gasfluxes through the elements and the desorption and the pumping speed. The calculations are matrix calculations in which the matrix contains the geometry of the system (volume, conductance and pumping speed). The matrix gives the relation between desorption "vector" and pressure distribution.
Consider the part of the system outlined in figure 1.

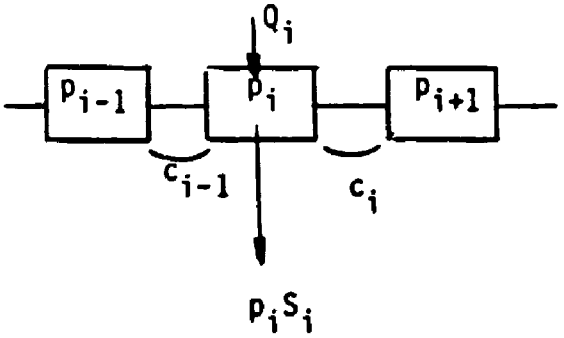

Fig. 1: Schematic set up of the system

Consider all i in the array of elements. When p_i, V_i, n_i and T_i are the symbols for respectively the pressure, the volume, particle density and temperature in cell i, the ideal-gas law can be written as: $p_i V_i = n_i k T_i$ (k = Boltzmann's constant). The desorption rate, Q_i in cell i is defined as the time derivative:

$$Q_i = \frac{d}{dt}(p_i V_i) = kT_i \frac{dn_i}{dt} = p_i \frac{dV_i}{dt} + V_i \frac{dp_i}{dt} = p_i S_i + V_i \frac{dp_i}{dt}$$

(isotherm) S_i is the pumping speed at cell i.

Furthermore the contributions to Q from neighbouring cells are given by $Q_{ji} = C_{ji}\ (p_j - p_i)$, where C_{ji} is the conductance between cell j and cell i.

(C_{ji} is calculated as $\frac{2C_iC_j}{C_i+C_j}$, where j=i-1 or j=i+1).

For element i we have:

$$Q_i + c_{i-1}(p_{i-1}-p_i) + c_i(p_{i+1}-p_i) - p_iS_i - V_i \frac{dp_i}{dt} = 0$$

To solve this system of differential equations the following approximations are made:

$$dp_i \simeq p_{it} - p_{it-\Delta t} \quad \text{and}$$

$$\ln(-Q_i) = \ln(-Q_{io})\ n_m - \ln(t-t_o)$$

{R.J. Elsey, Vacuum V25, 1975, p 302}

This results in the matrix relation:

$$A * \begin{pmatrix} p_{i_t} \\ p_{2_t} \\ \vdots \\ p_{N_t} \end{pmatrix} = \begin{pmatrix} Q_{0_1}.t^{-1} + \frac{V_1 p_{1_{t-\Delta t}}}{\Delta t} \\ Q_{0_2}.t^{-1} + \frac{V_2 p_{2_{t-\Delta t}}}{\Delta t} \\ \vdots \\ Q_{0_N}.t^{-1} + \frac{V_N p_{N_{t-\Delta t}}}{\Delta t} \end{pmatrix}$$

with

$$A = \begin{pmatrix} S_1 + C_1 + \frac{V_1}{\Delta t} - C_1 & 0 \ldots\ldots 0 \\ -C_1 & S_2 + C_2 + C_1 + \frac{V_2}{\Delta t} - C_2 \ldots\ldots \\ 0 & -C_2\ S_3 + C_3 + C_2 + \frac{V_3}{\Delta t} \end{pmatrix}$$

3. The vacuum system of the ESRP

In this chapter the method described in the previous chapters is applied to the design of the ESRP-case. Calculations of the pressure distributions are made for different pumpdown-times. The effect of the changes in the installed pumping speed on the pressure distribution is studied.
The ESRP machine is considered to be made up of 16 identical sections separated by sector valves. The calculations are carried out for one of those sections. A section is approximately 43 metres long and made up of dipole-, quadrupole- and sextupole magnet chambers, wigglers, undulators, pumpouts, valves and bellows.
In our program the section is divided in 52 elements (each component is one element) and for each element its volume, inner area, conductance and pumping speed are determined. In figure 2 an outline is given of one section.

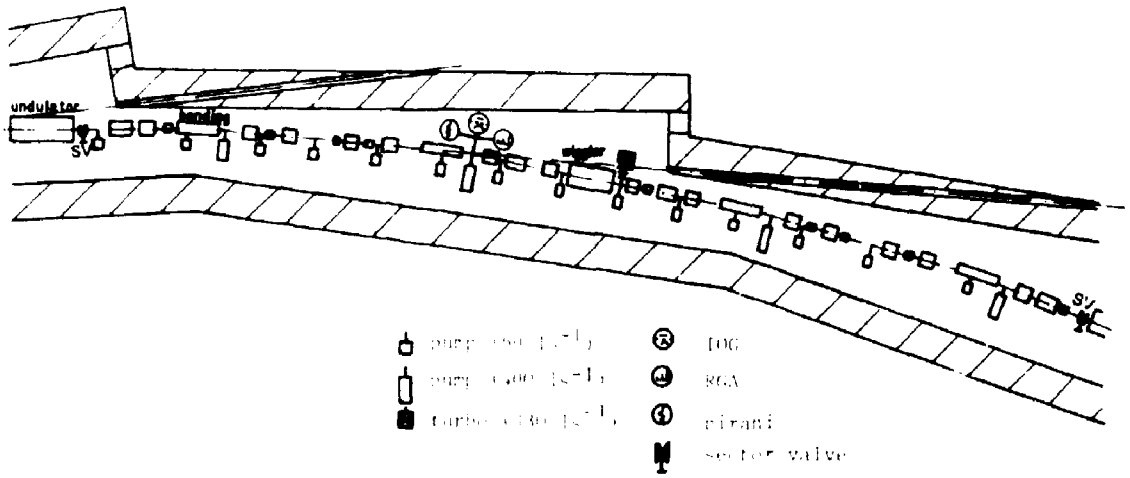

Fig. 2: 1/16 of the ESRP

The element-parameters are given in table 1.

Table 1.

NR		LENGTH (M)	VOLUME (L)	CONDUCTANCE (L/S)	PUMP SPEED (L/S)	DESORPTION (TR=L/S)
1	Sector valve	0.300	5.072	80.0	0.0	.15E-07
2	Pumpout	0.650	7.000	19.5	35.0	.11E-08
3	QF 7	1.800	1.192	5.9	0.0	.46E-08
4	QD 6	2.800	1.040	0.8	0.0	.40E-08
5	Bellow	2.900	0.503	165.0	0.0	.18E-08
6	S2+Dipole B1	5.608	11.170	3.0	35.0	.32E-07
7	Crotch system	5.920	38.637	31.0	198.0	.13E-06
8	Bellow	6.020	0.503	165.0	0.0	.18E-08
9	QD 4	7.170	1.192	5.9	0.0	.46E-08
10	Pump out	7.520	7.000	19.5	35.0	.48E-07
11	QF 5	8.720	1.244	5.7	0.0	.48E-08
12	Bellow	8.820	0.503	165.0	0.0	.18E-08
13	Pump out	9.170	7.000	19.5	35.0	.18E-02
14	SF + QF 5	11.290	2.229	3.2	0.0	.86E-08
15	Pump out	11.640	7.000	10.5	35.0	.11E-08
16	QD 4	12.840	1.244	5.7	0.0	.48E-08
17	Bellow	12.940	0.503	165.0	0.0	.18E-08
18	SD + Dipole B1	15.640	11.170	3.0	35.0	.32E-07
19	Crotch system	15.960	38.637	31.0	198.0	.13E-06
20	Bellow	16.060	0.503	165.0	0.0	.18E-08
21	QD 3	17.060	1.040	6.0	0.0	.46E-08
22	Pump out	17.410	7.000	19.5	35.0	.48E-07
23	QF 2	18.460	1.092	6.5	0.0	.42E-08
24	Bellow	18.560	0.503	165.0	0.0	.18E-08
25	QD 1	19.560	1.040	6.8	0.0	.40E-08
26	Pump out	19.910	7.000	19.5	35.0	.12E-07
27	Wiggler	22.160	2.332	3.0	0.0	.90E-08
28	Pump out	22.510	7.000	19.5	35.0	.11E-08
29	QD 1	23.310	0.829	8.5	0.0	.32E-08
30	QF 2+S 13	25.010	1.762	4.0	0.0	.68E-08
31	Pump out	25.360	7.000	19.5	35.0	.11E-08
32	QD 3	26.360	1.040	0.8	0.0	.40E-08
33	Bellow	26.460	0.503	165.0	0.0	.18E-08
34	Dipole B1	29.168	11.170	3.0	35.0	.32E-07
35	SD + Crotch syst.	29.480	38.637	31.0	198.0	.13E-08
36	Bellow	29.500	0.503	165.0	0.0	.18E-08
37	QD 4	30.730	1.192	5.9	0.0	.46E-08
38	Pump out	31.080	7.000	19.5	35.0	.48E-07
39	QD 5+S 18	32.730	1.710	4.1	0.0	.66E-08
40	Bellow	32.830	0.503	165.0	0.0	.18E-08
41	Pump out	33.280	7.000	15.1	35.0	.12E-07
42	QF 5	34.938	1.710	4.1	0.0	.66E-08
43	QD 4	36.500	1.710	4.1	0.0	.66E-08
44	Bellow	36.680	0.503	165.0	0.0	.18E-08
45	SD + Dipole B1	39.388	11.170	3.0	35.0	.32E-07
46	Crotch system	39.708	38.637	31.0	198.0	.13E-06
47	Bellow	39.800	0.503	165.0	0.0	.18E-08
48	QD 6	40.850	1.088	6.5	0.0	.42E-08
49	Pump out	41.200	7.000	19.5	35.0	.48E-07
50	QF 7+S 22	42.550	1.408	5.0	0.8	.54E-08
51	Pump out	42.900	7.000	19.5	35.0	.12E-07
52	Section valve	43.200	5.072	80.0	0.0	.15E-07

Condition nr = .29E+02

Starting from the data in table 1, the pressure distributions as a function of time in the section is calculated and given in figure 3.

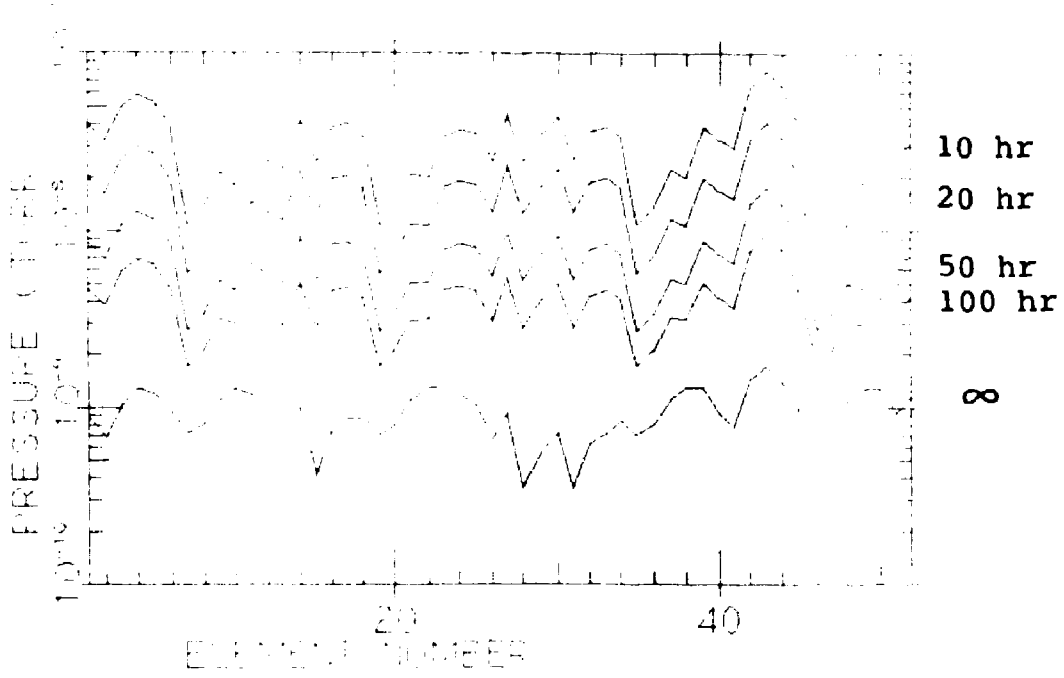

Fig 3: Pressure contributions as a function of time.

The system is pumped by sputter-ion pumps (S = 60 l/s and S = 400 l/s). The system is pumped down from atmospherical pressure to a pressure < 10^{-5} Torr (starting pressure of the sputter-ion pumps) with a turbo molecular pump. One turbo molecular pump (S = 330 l/s) is used per section. The (dynamical) pressure distribution of the pumpdown to 10^{-5} Torr is given in figure 4.

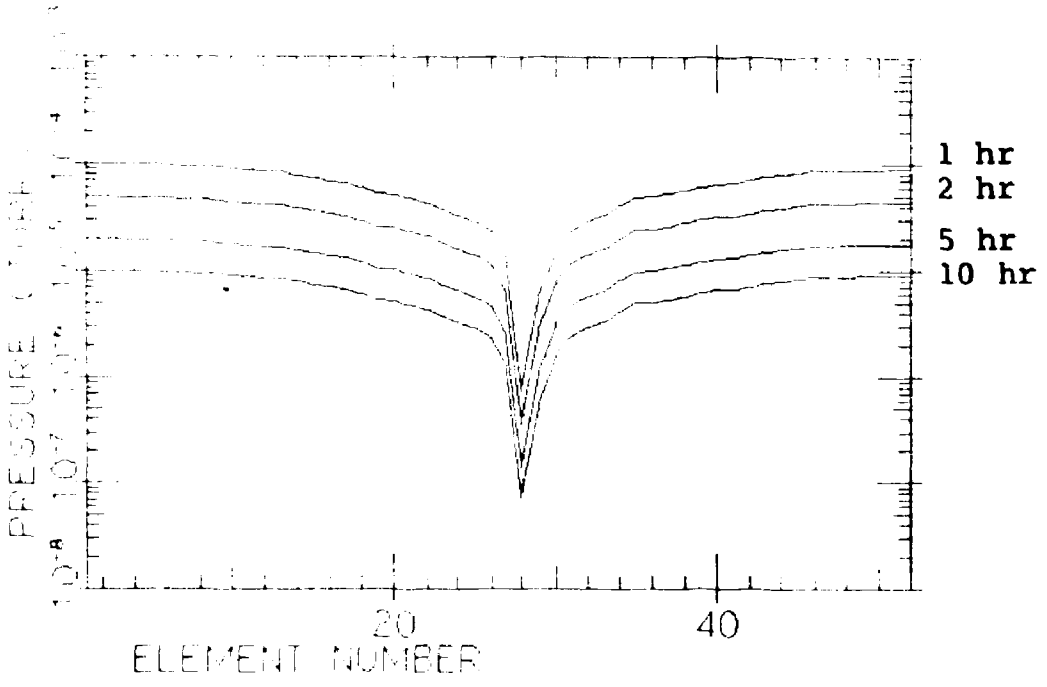

Fig. 4: Pressure distribution during pump-down

3.1 Design considerations

In general integrated ion-pumps are used in the dipole chambers of storage rings. From the calculations it is clear that an average endpressure of 10^{-9} Torr can be achieved by using distributed lumped sputter-ion pumps. There are two main reasons why no integrated ion pumps are applied in this design study:

1. Mostly the magnetic field of the dipole magnets is used for the operation of integrated ion-pumps. The bending magnets of the ESRP-machine have a relatively low magnetic field strength and for this reason special attention has to be paid to the functioning of integrated ion pumps.
2. In general the synchrotron radiation is absorbed in the dipole chamber, resulting in a more or less uniform desorption in the dipole chambers.

In this ESRP-design the synchrotron radiation is absorbed at defined places (crotches). The desorption caused by the synchrotron radiation can be removed almost (> 90%) completely by installing a large pumping speed at the crotch (S_{eff} = 400 l/s).

3.2 Total installed pumping speed

Sixteen sputter-ion pumps (S_{eff} = 35 l/s) and 4 sputter-ion pumps (S_{eff} = 198 l/s) together with one turbo-molecular pump (S = 330 l/s) are installed in one section. In section 4.2.1 a study is made of the effective total speed for several values of the total pumping speed. In section 4.2.2 we study the effect of malfunctioning of the sputter-ion pumps. This is important in the design we suggest that several (f.i. 4) sputter-ion pumps are connected to the same power supply.

3.2.1 Efficiency of the pumping speed

The pressure distribution and average pressure are determined for a pumping speed ten times larger and ten times smaller than the design value (S_{design}= 1400 l/s). See figure 5.

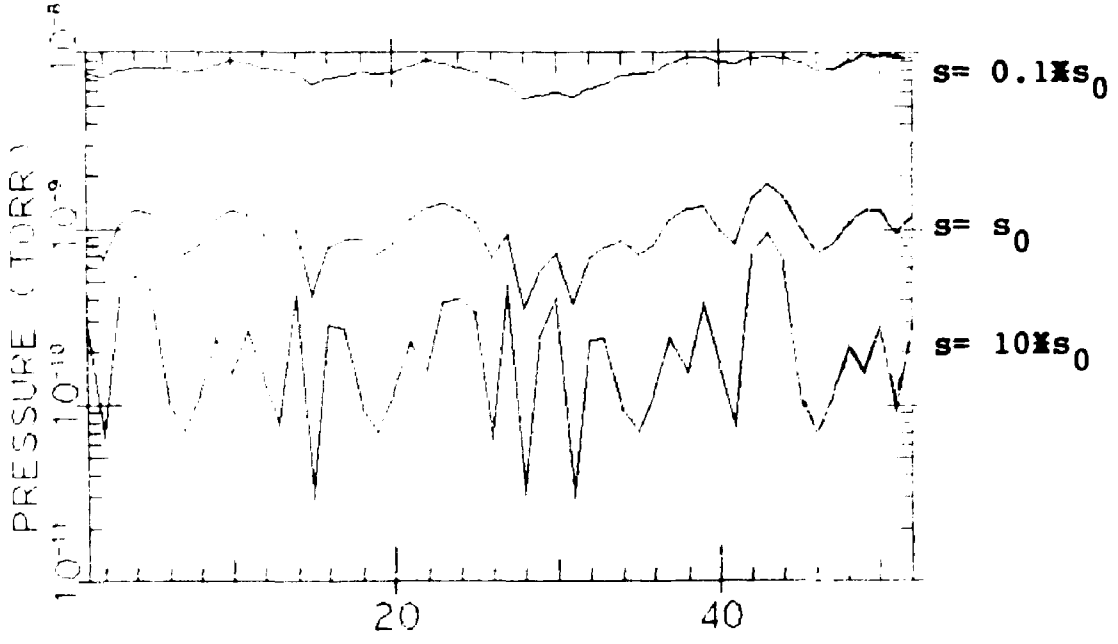

Fig. 5: Influence of pumping speed on the pressure distribution.

Table 2 gives the total pumping speed, the average pressure, the effective pumping speed and the efficiency of the pumping speed.
The total desorption of a section is 0.1×10^{-5} Torr l/s.

S installed (l/s)	p_{ave} (Torr)	S_{eff} (l/s)	Efficiency (%)
0.14×10^{3}	0.80×10^{-8}	0.14×10^{3}	98
0.14×10^{4}	0.10×10^{-8}	0.11×10^{4}	79
0.14×10^{5}	0.35×10^{-9}	0.31×10^{4}	22

Table 2: Efficiency of total pumping speed.

The effeciency being $\frac{S_{eff}}{S_{tot}} \times 100\%$ of the turbo-molecular pump is also studied.
The pressure distribution is determined for a turbo-molecular pump of 33. 330 and 3300 l/s pumping speed. See figure 6.

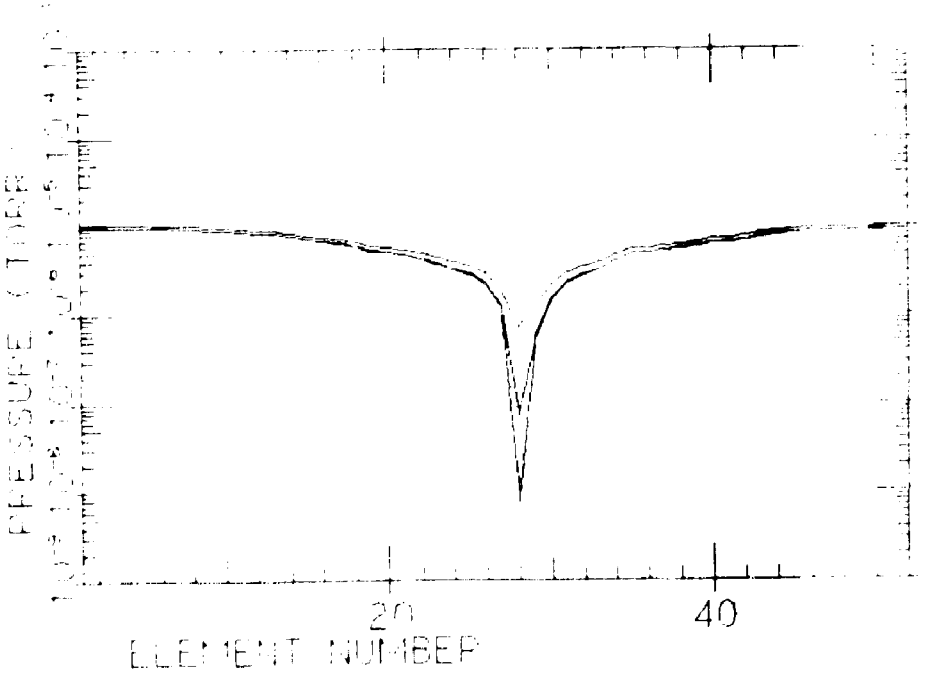

Fig. 6: Efficiency of the pump-down unit

Table 3 gives the pumping speed, average prrssure, effective pumping speed and the efficiency.

S (l/s)	p_{ave} (Torr)	S_{eff} (l/s)	Efficiency (%)
33	0.10×10^{-4}	0.11	0.33
330	0.10×10^{-4}	0.11	0.33×10^{-1}
3300	0.10×10^{-4}	0.11	0.33×10^{-2}

3.2.2 Malfunctioning of the sputter-ion pumps (S_{eff} = 35 l/s)

Malfunctioning of half of the sputter-ion pumps a its effect on the end-pressure is studied in this chapter. The pressure distribution of 1/16 ESRP-section pumped with 8 sputter-ion pumps (S_{eff} = 35 l/s) and 4 sputter-ion pumps (S_{eff} = 198 l/s) is given in figure 8.

The pumps are connected in the following way (figure 7) to the power supplies

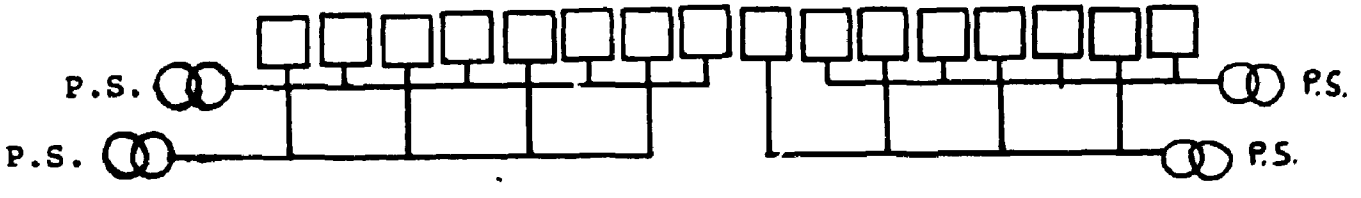

Fig. 7: The way of sharing power supplies

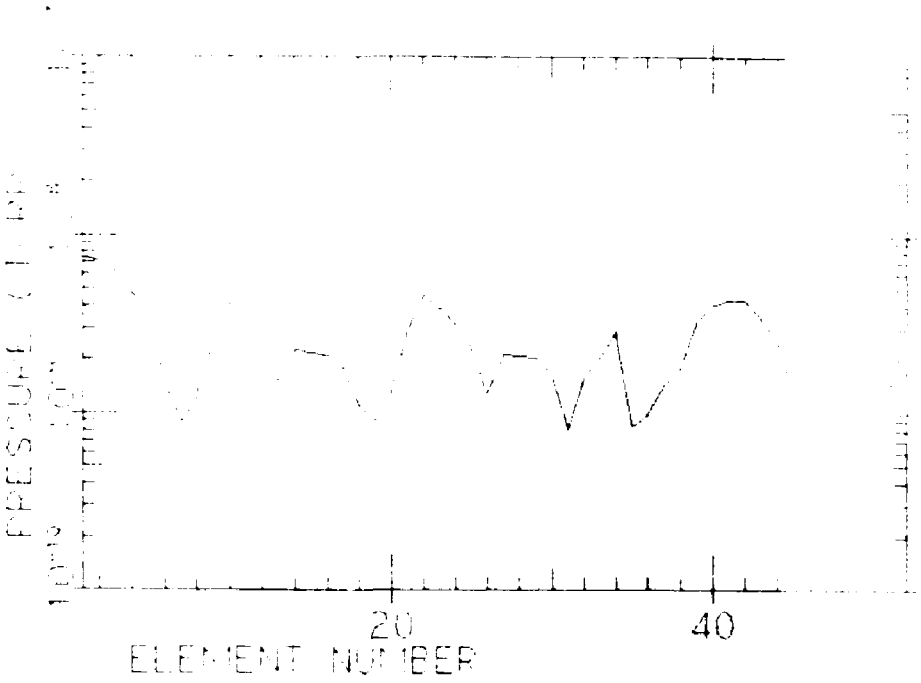

Fig. 8:

We suggest special cables are used to connect the pumps to the power supplies in order to power the other pumps in case of a short-circuit in one pump. (ref. Dr. Poncet)

3.2.3 Conclusions

As can be seen in table 3 the efficiency of pumping speed of the design is 79%. When 50% of the sputter-ion pumps (S = 60 l/s) is switched off the average pressure will increase to 6×10^{-9} Torr resulting in a beam life-time of about 5 hours. The effective pumping speed for the turbo-molecular pump is determined by the conductance of the section and not by the pumping speed of the pump.

3.3 Influence of stepsizes

In this section the influence of the size of the elements and the timestep Δt on the calculated pressure distributions will be studied.

In the calculation of the pressure in an element it is assumed that desorption and pumping speed act in one point: the nodal point.
The pressure is considered to be constant in the element. For an exact calculation of the pressure distribution the elements have to be chosen of infinitesemal size, because then pressure, desorption and pumping speed can considered to be constant. To get an approximation the first quarter of the section, -13 elements- wer divided into 10 parts (the dipole chamber into 30 parts). Because of shortage of computer-memory (PDP10) it was not possible to carry out the calculations for the whole section.

Figure 9 gives the pressure distribution for the first 13 elements of the section.

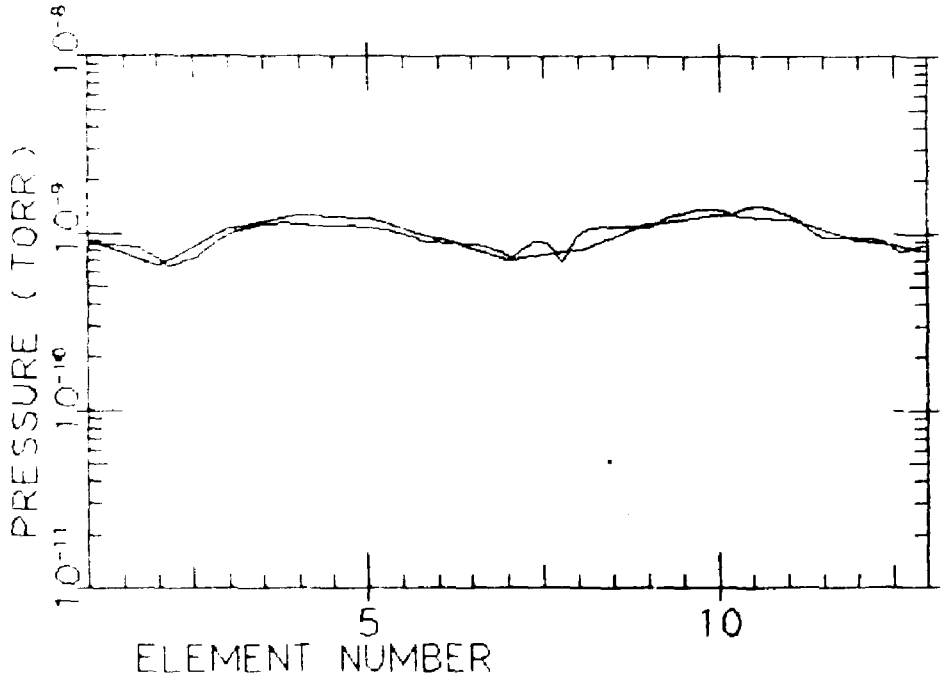

Fig. 9: The influence of the stepsize.

In the dynamical calculations a time-step Δt has to be introduced. It is important to know the influence of the size of the time-step on the calculated pressure distribution. The pressure distribution after 3600 s. is calculated for a time-step of 720, 360 and 180 s.
As can be seen the influence of the time-step is negligable.
The result showed a difference of less than 3%.

3.4 Conclusions

Calculations indicate a total of 256 starcell-type of ion pumps (S = 60 l/s^{-1}) and 64 conventional sputter-ion pumps (S = 400 ls^{-1}).
An extra blindflange on those pumps provides for the installation of multifilament titanium sublimation pumps as an option. Total desorption of ESRP during operation is estimated on 1.8×10^{-5} Torr ls^{-1} and a total (eff.) pumping speed of 2.24×10^{4} ls^{-1} (distributed in the proposed way) results in an average pressure of 1.1×10^{-9} Torr (79% effective pumping speed). Dynamical calculations indicate a pumpdown time of about 500 h.
A total of 16 turbo-pumps is used as roughing system (one per sector).
Precaution has to be taken to eliminate the possibility of oil contamination. The turbo-molecular pumps are connected to the ring via all metal 4" valves. There are no valves between the sputter-ion pumps and the ring. The ring vacuum is monitored by the current of the sputter-ion pumps.
Each sector will be equipped with a pirani head, one IOG and a quadrupole residual gas analyser. The vacuum measuring equipment has to be coupled to the computer.

4. Cost estimates

In this section we have summarized the costs of the mechanical vacuum equipment for the storage ring itself. Most of the prices are based on quotations, although competetive firms have not yet been invited to quote.

Sharing power supplies and control units can save about 190 of them. This cost saving is dependent on the need for local pressure information.
The budget is based on the situation where 4 sputter-ion pumps are connected to one power supply in such a way that in between each two pumps one of the next set of pumps is located (fig. 7) (in order to maintain a reasonable pressure distribution if one power supply is out of order). The 400 l/s^{-1} pumps are not combined.
The pressure measurement is mainly based on the indication of the 400 l/s pump power supply and the combined control units of the starr cells.
In order to have a more precise measurement 16 ionisation gauges are implemented which also serve as a trigger for closing the sector valves.
The piranies give the vacuum information in the high pressure region (f.e. pumpdown).

Quadrupole mass spectrometers are used to analyse the system during pumping in the UHV region. From these spectra one can decide of baking is necessary or a leak is limiting the ultimate pressure.
The all metal sector valves are needed in order to separate 1/16th of the circumference of the ring and to isolate the 4 cavities from the rest of the system.
Five crotches per straight section are sufficient to absorb the main beam power.
The bake-out equipment consists of heating tapes and rings, isolation material and a cooled copper shielding inside the dipoles.

The total budget for bake-out includes an estimate of the control equipment needed to regulate the power as a function of location and time. The price for the fabrication of the 4 cavities is not estimated.
Drafting and engineering will take 32 my while pre-assembly and tests are estimate to take another 32 my.
Installation needs 20 my.
Building prototype vessels, cavities and crothes and finalise the design will cost at least one year of intensive study.

A. Pumps

	Item	Number	Price each	Total x 1000 Sfr
I.	60 l/s^{-1}	260		702,5
	bake out mantle	260		39
	power supplies	68		107,7
	control units	68		117,7
II.	400 l/s^{-1}	76 (cavities 12)		858,9
	bake out mantle	76		
	power supplies	76		
III.	turbo units	18		443,5
			Subtotal	2269,3

B. Pressure measurement

Item	Number	Price each	Total x 1000 Sfr
ionisation gauges IMR 125 + IMG 070	32	7274	232,8
pirani's TPG 035	32	2641	84,5
quadrupole m.s. QMG 112	16	24800	396,1
		Subtotal	713,4

C. Valves

Item	Number	Price each	Total x 1000 Sfr
isolation valves	26(16+8+2)	18436	479,3
evacuation valves	21(16+4+1)	20894	438,7
venting valves	18(16+2)	7374	132,7
		Subtotal	1050,8

D. Mechanics

Item	Number	Price each	Total x 1000 Sfr
crotches	320	1000	320
dipole vessel	64	35000	2240
bake out	64	12000/straight	768
bellows	256		160
"triplet" vessel	64	20000	1280
monitors	256	4000	256
flanges	1600		299
		Subtotal	5323,0
	Total		9356,5

Excluded: Tax, overhead, cavities, magnets, design, cleaning facilities, testing facilities, cabling, tools and allignment.

5. Appendix

In this chapter the mathematical framework of the calculations will be discussed. The condition number of the matrix is used to characterize the sensitivity to errors for a system of linear equations. Two methods are given for the determination of the least square solution of an mxn (m > n) system of linear equations (singular value decomposition and normal equations of Gauss).

5.1 Condition of a matrix

The elements of a matrix A and a right-handside vector $\vec{b}$ are rarely precisely known (measurement errors, round-off errors in the floating point processor). It is important to know how disturbances in A and b influence the solution of the system $A\vec{x} = \vec{b}$. A matrix with two identical rows or columns is singular. A singular system has no or an infinite number of solutions. In a nearly singular system small changes in A or $\vec{b}$ will cause significant changes in the solution. The condition number of a matrix is a measure for the singularity of that matrix. Before the condition number is discussed the vectornorm has to be introduced. Throughout this paper we use the Euclidian norm:

$$//\vec{x}//_2 = (\sum_{i=1} /\vec{x}_i/^2)^{\frac{1}{2}}$$

When a vector $\vec{x}$ is multiplied by a matrix A we get a new vector $A\vec{x}$, which in general has a different norm than $\vec{x}$. The changes in this norm are directly related to the sensitivity to errors in A and $\vec{b}$. The range of possible changes are characterized by two numbers: m and M.

$$M = \max_{\vec{x}} \frac{//A\vec{x}//}{//\vec{x}//} \qquad m = \min_{\vec{x}} \frac{//A\vec{x}//}{//\vec{x}//}$$

The condition number is defined as:

$$\text{cond}(A) = \frac{M}{m} = \frac{\max\limits_{\vec{x}} \frac{//A\vec{x}//}{//\vec{x}//}}{\min\limits_{\vec{x}} \frac{//A\vec{x}//}{//\vec{x}//}}$$

when $m = 0$ the matrix A is singular.
Suppose: $A\vec{x} = \vec{b}$ and $A(\vec{x} + \Delta\vec{x}) = (\vec{b} + \Delta\vec{b})$ so $A\Delta\vec{x} = \Delta\vec{b}$
then $//\vec{b}// \leq M//\vec{x}//$ and $//\Delta\vec{b}// \geq m//\Delta\vec{x}//$

when $m \neq 0$ $\dfrac{//\Delta\vec{x}//}{//\vec{x}//} \leq \text{cond}(A) \dfrac{//\Delta b//}{//\vec{b}//}$

The condition number is an upper limit for the magnification of the relative error.

5.2 QR and SVD algorithms

When using Gauss method or LU-decomposition the system $A\vec{x} = \vec{b}$ is multiplied by special matrices $p_j (j = 1,N)$ in order to get a system $A^{(n)}\vec{x} = \vec{b}^{(n)}$ which can be easily solved. As discussed in the previous chapter the sensitivity of $\vec{x}$ for errors in $A^{(j)}$ and $\vec{b}^{(j)}$ is determined by $\text{cond}(A^{(j)})$. When $\varepsilon^{(j)}$ is the (round-off) error caused by the transformation of $\{A^{(j-1)}, \vec{b}^{(j-1)}\}$ to $\{A^{(j)}, \vec{b}^{(j)}\}$, then this error can show up in $\vec{x}$ magnified by a factor $\text{cond}(A^{(j)})$.

So $\dfrac{//\Delta\vec{x}//}{//\vec{x}//} \leq \sum_{j=0}^{n-1} \varepsilon^{(j)} \text{cond}(A^{(j)})$

It will be clear that matrices p_j have to be selected in such a way as not to increase $\text{cond}(A^{(j)})$.
When U is an orthonormal matrix then $U^TU = I$
$//A// = //U^TUA// = //U^T//.//UA//$ and the
norm of an orthonormal matrix is equal to 1. $\}$ $\rightarrow$ $//A// = //UA//$.
So when we take p_j an arthonormal matrix than the condition of $(A^{(j)}\vec{x} = \vec{b}^{(j)})$ will not be influenced.

Furthermore p_j has to be chosen in such a way to simplify the matrix $A^{(j)}$. In general this can be written as:
$p.A = R$ and $p.\vec{b} = \vec{b}^*$ with p = orthonormal matrix and R = upper triangular matrix.
From this we get $A = p^T Q$ $\quad p^T = p^{-1} = Q$

$$A = QR \qquad \vec{b}^* = Q\vec{b} \quad .$$

Q and R can be calculated using Householder transformations.
The upper triangular matrix R can be written as a product of a diagonal matrix S and an orthonormal matrix V^T: $R = SV^T$, $s_{ij} = 0$ $(j \neq i)$.
An mxn matrix A can be written as: $A = QSV^T$ with Q = mxm matrix, S = mxn matrix and V^T= nxn matrix.
Because of the zeros, under the diagonal S we can drop the corresponding columns in Q changing Q to a mxm orthonormal matrix U.
We get $A = USV^T$.
Using singular value decomposition (SVD) the elements of S are chosen to be positive: the elements are called the singular values of A.
U contains the singular vectors of A.

5.3 Last square solution of a system of linear equations

A system of m equations with n variables (m >n) is often inconsistent. There has to be searched for a solution which is the best fit for the system of equations. The least square solution of an overdetermined system is given as the solution of the normal equations of Gauss:
$A^T A\vec{x} = A^T\vec{b}$.
The solution is $\vec{x} = (A^T A)^{-1} A^T\vec{b}$ with $A^+ = (A^T A)^{-1}A^T$ is called the pseudo- or Moore Penrose inverse.
In the SVD method the pseudo-inverse is written as
$A^+ = VS^{-1}U^T$.

5.4 Eigenvalues/vectors and singular values/vectors

An eigenvector $\vec{e}$ of a square matrix A is a vector ($\vec{e} \neq \vec{0}$) which satisfies the relation $A\vec{e} = \lambda\vec{e}$ ($\lambda, e\varepsilon$). λ is called the eigenvalue of A. For an mxn matrix the singular values/vectors take the place of the eigenvalues/vectors. The singular values of A are the positive square roots of the eigenvalues of A^TA.

In section 2.1 we have: $\text{cond}(A) = \frac{M}{m} = \dfrac{\max\limits_{\vec{x}} \frac{//A\vec{x}//}{//\vec{x}//}}{\min\limits_{\vec{x}} \frac{//A\vec{x}//}{//\vec{x}//}}$

$$//\lambda\vec{x}// = /\lambda/.//\vec{x}// \leq M//\vec{x}// \rightarrow M \geq \lambda$$

$$/\lambda/.//\vec{x}// \geq m//\vec{x}// \rightarrow m \leq \lambda$$

When λ_{max} is the (absolute value) largest eigenvalue and λ_{min} the smallest we get:

$$\text{cond}(A) = \frac{M}{m} \geq \frac{\lambda_{max}}{\lambda_{min}}$$

For an overdetermined system we get: $\text{cond}(A) \geq \frac{\sigma_{max}}{\sigma_{min}}$.

5.5 Error-analysis using eigenvector-decomposition

This analysis can be used to determine to which degree, in a system $A\vec{x} = \vec{b}$, a disturbance in $\vec{b}$ will influence the solution $\vec{x}$.

Suppose $A(\vec{x} + \Delta\vec{x}) = \vec{b} + \Delta\vec{b}$; $A\Delta\vec{x} = \Delta\vec{b}$ (A = nxn matrix).

For an eigenvector $\vec{e}_i$ of A we have: $A\vec{e}_i = \lambda_i\vec{e}_i$. When Δb is located in the direction of $\vec{e}_i$ then $\vec{x}$ will be influenced by a factor $\frac{\Delta b}{\lambda_i}$.

In general $\Delta\vec{b}$ can be decomposed in the basis of eigenvectors:

$$\Delta\vec{b} = c_1\vec{e}_1 + c_2\vec{e}_2 + \ldots. c_n\vec{e}_n$$

A measure of the error-amplification in direction $\vec{e}_i$ can be written as c_i/λ_i.

Using SVD (useful for mxn matrices) the vector $\Delta\vec{b}$ can be written as a linear combination of the singular vectors $\vec{U}_i$ of A:

$$\Delta\vec{b} = c_1\vec{U}_1 + c_2\vec{U}_2 + \ldots. c_n\vec{U}_n.$$

The value $\frac{c_i}{\sigma_i}$ (σ_i = singular value of A) is a measure of the erroramplification in direction $\vec{U}_i$.