UNIVERSITÀ DEGLI STUDI DI PAVIA

# SCIENTIFICA ACTA

## QUADERNI DEL DOTTORATO

VOLUME XX ANNO XX NUMERO 4

15 DICEMBRE 2005


ISTITUTO NAZIONALE DI FISICA NUCLEARE

CENTRO STAMPA – DIPARTIMENTI FISICI
Via Bassi, 6 27100 Pavia

# INDEX

# Polymers and Alcohols: Soft-Matter problems and the statistical mechanics of chains

Marco Pieruccini
CNR - Istituto per i Processi Chimico-Fisici (Sez. Messina)
Via La Farina 237, 98123 Messina, Italy.
email: pieruccini@me.cnr.it

April 18, 2005

# 1 A reminder of statistical thermodynamics

*"Thermodynamics is quite different [from mechanics and electrodynamics]. It neither claims a unique domain of systems over which it asserts primacy, nor does it introduce a new fundamental law analogous to Newton's or Maxwell's equations. In contrast to the specificity of mechanics and electromagnetism, the hallmark of thermodynamics is generality. Generality first in the sense that thermodynamics applies to all types of systems in macroscopic aggregation, and second in the sense that thermodynamics does not predict specific numerical values for observable quantities. Instead, thermodynamics sets limits (in- equalities) on permissible physical processes, and it establishes relationships among apparently unrelated properties."* (H.B. Callen: "Thermodynamics and an introduction to thermostatistics")

These sharp features, so nicely outlined by Callen, are at the basis of the fundamental role that thermodynamics plays in approaching the study of macroscopic systems in general, and in particular of those which constitute the rather wide class of the so called Soft Matter. Biological systems, polymers and associating liquids are among the systems of interest in this context. Their complexity, and with it the need of an interdisciplinary approach to the investigation of their properties, takes us away from the position where

*"... [physicists] have the habit of taking the simplest example of any phenomenon and calling it 'Physics', leaving the more complicated examples to become the concern of other fields."* (R. Feynman, *The Feynman Lectures on Physics*).

It is in the spirit of keeping as close as possible to this ideal situation that thermodynamics exerts its attractiveness; but it is of course its force to establish its relevance in this field. Further, coupling to the microscopic world through the statistical mechan-

ical foundation of thermodynamics, represents as a consequence a particularly fruitful strategy of investigation.

In the following, the most important classical thermodynamical concepts relevant to the main body of these lectures, namely, the thermodynamic potentials, will be briefly recalled. The statistical mechanical counterparts, i.e. the appropriate partition functions, will be considered afterwards. There are a number of textbooks which can be taken for a reference for these arguments; among the others the reader may consult refs. [1].

## 1.1 Thermodynamic potentials

### 1.1.1 Survey of general principles

The equilibrium thermodynamic state of any macroscopic system is completely described by an extensive function

$$S \equiv S(E; \mathbf{X}), \tag{1}$$

the entropy, which depends on the energy $E$ of the system and on other extensive variables $\mathbf{X}$ (i.e. volume $V$, particle number $N$, total electric moment $\mathbf{M}$, etc.); the derivative

$$\frac{1}{T} \equiv \left(\frac{\partial S}{\partial E}\right)_{\mathbf{X}} \tag{2}$$

is always positive, and defines the temperature of the system.

The ***first law*** of thermodynamics expresses the conservation of the energy: the overall energy content $U$ of a macroscopic system whatsoever may change only as a consequence of either a work $\delta W$ performed by (or on) the system, or an inwards/outwars heat flux $\delta Q$ through the boundaries, or else because of a change $\delta N$ in the number of particles of the system:

$$dU = \delta Q - \delta W + \mu\, dN\,. \tag{3}$$

The intensive parameter $\mu$ is the "chemical potential", and carries with it not only interaction energy with other particles, but thermal

energy as well (in the case of a perfect gas, in fact, only this component is present). This quantity is very important for the description of phase equilibria and flows of matter.

Work is a means by which energy can flow in or out of a system. The heat is also a form of energy flux; it *measures* the entropy change too, but only when the heat transfer process is performed reversibly (see below); in this case

$$dS = \frac{\delta Q_{rev}}{T} \,. \tag{4}$$

The ***second law*** of thermodynamics concerns the spontaneous changes of $S$ when internal constraints within an isolated system are released. In general, by constraint is meant any means preventing a spontaneous transition of a system towards a different thermodynamic state. So, for instance, an adjustable piston separating two chambers into a cylinder filled by a gas, as well as a means whatsoever able to set locally an arbitrary chain orientation in a polymer melt, are both examples of internal constraints.

Now, consider a system initially in equilibrium, that is, with a well established $S(E, \mathbf{X})$. Then, suppose to apply an internal constraint in order to drive the system towards a different state with entropy $S'(E, \mathbf{X}; \text{constr.})$, the other extensive parameters remaining fixed. After this new state has been reached, isolate the system completely and release the constraint. The system will spontaneously return to the initial state, and the second law states that the entropy change associated to this transformation is always positive, i.e.

$$S - S' > 0 \,. \tag{5}$$

This general relationship underlies also the transformations involving any non-isolated macroscopic system; indeed we shall always consider the latter as a small part of a much larger, globally isolated ("total") system to which eq. 5 applies. The rest, will be the "environment" of the non-isolated (sub-) system. We take this consideration as the route (among the others) to introduce the thermodynamic potentials, i.e. those quantities whose variation

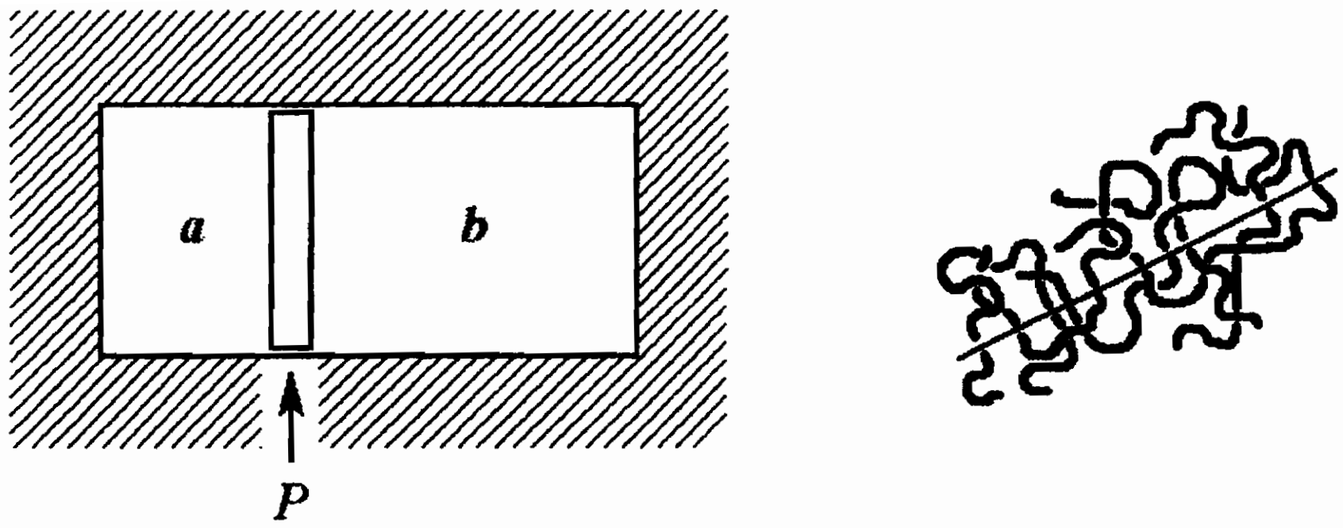





Figure 1: A fixed piston **P**, separating two chambers **a** and **b** filled by a gas at different pressures but same temperature, is an example of internal constraint. If **P** is released, then the system will change spontaneously towards a state of higher overall entropy (the environment included). A mechanism whatsoever able to force the local chain orientation along a given direction in a polymer melt represents an internal constraint too.

accompanies the thermodynamic transformations in analogy with the mechanical behavior of a conservative system.

From the additivity property of the entropy we can always write

$$TdS_{tot} = TdS + TdS_{env}\,, \tag{6}$$

where from now on, quantities without the subscript refer to our system, while the subscripts *"tot"* and *"env"* are put explicitly to refer either to the total system, or to the part of it which is taken as environment.

Let $\delta Q_{env}$ be the heat to be transferred reversibly to the environment in order to determine a change $dS_{env}$, and let this heat be supplied by the system, so that $\delta Q_{env} = -\delta Q$. We thus may write eq. 6 in terms of quantities entirely relating to the system as

$$TdS_{tot} = TdS - \delta Q\,. \tag{7}$$

We assume that the environment behaves as a reservoir, and that the only spontaneous processes take place within the system; then,

eq. 7 holds in general for both reversible and irreversible transformations occurring in the system. In particular, it relates any entropy change of the *total* system to the degree of irreversibility of the process occurring in the (sub-) system, i.e. to the difference between $dS$ and $\delta Q/T$. By resorting to the first principle we can express eq. 7 in terms of the work as

$$T dS_{tot} = T dS - dU + \mu\, dN - \delta W\,. \tag{8}$$

From eq. 5 we know that the r.h.s. of eq. 8 cannot be negative, and it is zero only if the process is reversible, i.e. if the representative point of the whole isolated system moves on a globally iso-entropic hypersurface.

Equations 5 and 8 are the fundamental relations upon which the concept of thermodynamic potential find its basis. Before developing this argument, however, we will first introduce some further basic property relating to the functional dependence of the entropy on the other relevant thermodynamic variables.

### 1.1.2 Euler's equation and Gibbs-Duhem relation

In the fundamental relation of thermodynamics, eq. 1, both the entropy $S$ and its arguments are extensive variables, so the following relation holds (Euler' equation)

$$S = \frac{1}{T} U + \boldsymbol{\xi} \cdot \mathbf{X}\,, \tag{9}$$

where $\boldsymbol{\xi} \equiv \partial S/\partial \mathbf{X}$ are the *intensive* variables conjugated to $\mathbf{X}$ (for instance, $\xi = p/T$ if $\mathrm{X} \equiv V$, $\xi = -\mu/T$ if $\mathrm{X} \equiv N$). Note that eq. 9 is a sharp equality.

If we take the total differential of last equation and recall the first law, eq. 3, then we find the Gibbs-Duhem equation

$$S\, dT + \mathbf{X} \cdot d\boldsymbol{\xi} = 0\,, \tag{10}$$

which establishes a relationship among the intensive variables of a system in thermal equilibrium. This is a very useful tool to find the actual values of an intensive variable once the others are known.

### 1.1.3 Helmholtz potential

We now turn back to the problem of finding the thermodynamic functions which describe the direction of the spontaneous processes in terms of the variables of the system.

Define the Helmholtz free energy for a system with given volume, number of particles and temperature as

$$F \equiv U - TS\,; \tag{11}$$

then, for an isothermal process eq. 8 reads

$$\delta W = -dF - T dS_{tot}\,, \tag{12}$$

where $\delta W$ stands for the work done by the system. Since $dS_{tot}$ cannot be negative, last equation means that the maximum work can be performed only as a result of a reversible transformation (i.e. when $dS_{tot} = 0$), and in this case it equals the change of the Helmholtz free energy (potential), much like in the case of a conservative mechanical system.

If no work is allowed to be done, i.e. $\delta W = 0$, irreversibility (the spontaneous settlement into an equilibrium state, i.e. with $dS_{tot} > 0$ sharp) implies that $F$ changes in order to reach a minimum. Equation 12 generalizes the concept of a conservative mechanical system in that, together with work, it includes also the flux of *non-organized* energy, i.e. the heat.

### 1.1.4 Gibbs potential

Consider a system able to change its volume (but not matter) at a constant pressure $p$ during a transformation whatsoever, and let $\delta W_0 \equiv \delta W - p\,dV$ be the work other than $p\,dV$ performed; then, from eq. 8 this work is given by

$$\delta W_0 = -dG - T dS_{tot}\,, \tag{13}$$

where the quantity

$$G \equiv U + pV - TS = N\mu \tag{14}$$

is the Gibbs free energy (last equality in eq. 14 is a consequence of the Euler's equation). As before, $|\delta W_0|$ is a maximum when the process is reversible, and if no $\delta W_0$ is allowed, then any possible spontaneous evolution of the system will drive it towards a state of minimum $G$ (we stress that this circumstance is a direct consequence of eq. 5).

### 1.1.5 Enthalpy

From eqs. 6, 13 and 14 we derive immediately

$$\delta W_0 = -dH - T dS_{env} \,, \tag{15}$$

where the quantity $H \equiv U + p\,dV$ is called the enthalpy of the system (in the process described by eq. 15 no exchange of matter is allowed). Then, in the case that no heat is exchanged with the environment the enthalpy assumes the role of a potential for the work $W_0$ other than $p\,dV$.

### 1.1.6 Grand potential

This case concerns with the situation where a flux of matter between the system and the environment is possible, the volume remaining a constant. This can be accounted for through the term in eq. 3 describing the *energy* change (i.e. including heat) associated to the variation of the particle number. Then, by substitution into eq. 8 and by the same procedure leading to eq. 12, we find:

$$\delta W = -d\Phi - T dS_{tot} \,, \tag{16}$$

where the Grand potential is defined by

$$\Phi \equiv U - TS - \mu N \,, \tag{17}$$

and it shares with $F$ and $G$ all the relevant features with regards to spontaneous transitions and maximum work.

### 1.1.7 Summary

In conclusion, (minus) the entropy is a potential for an isolated system, where all extensive variables are held fixed. The Helmholtz free energy $F$ is a potential for a system where the volume is held fixed together with particle number $N$ and temperature $T$, but its energy $U$ (i.e. the extensive variable conjugate of $T$) is free to fluctuate. The Gibbs free energy $G$ is a potential when $N$ is fixed together with $T$ and $p$, but the extensive variables $U$ and $V$, conjugated to the latter are let free to change. Finally, the grand potential describes the system when the volume $V$ is fixed, together with $T$ and $\mu$, but the conjugated extensive variables $U$ and $N$ are let free to change.

## 1.2 Statistical thermodynamics and fluctuations

### 1.2.1 The Boltzmann relation

The central relationship connecting thermodynamics to the microscopic states of a system made of a huge number of molecules is represented by the Boltzmann formula

$$S_{tot} = k_B \ln \Omega_{tot}(E_{tot}, \mathbf{X}_{tot}) \,. \tag{18}$$

This equation links the entropy of an isolated system in thermal equilibrium, with the total number $\Omega_{tot}$ of microstates compatible with the given (constant) overall values $E_{tot}$ and $\mathbf{X}_{tot}$ (a collection of identical systems in these conditions is called a *microcanonical ensemble*).

The additivity of the entropy for non-interacting (in fact *weakly* interacting) systems is verified by eq. 18, since the total number of microstates of the composite system equals the product of the number of microstates of each component. The above formula is also able to describe the effect of internal constraints; the latter reduce the number of microstates accessible to the system, and when they are released then the entropy increases, in agreement with the second law of thermodynamics (eq. 5).

The reduction of $\Omega_{tot}$ as an effect of the internal constraints should be considered in some detail because it is connected with the glass transition process. The idea underlying eq. 18 is about a restlessly moving representative point in the phase space of the system. When an internal constraint sets in, not necessarily *all* microstates with the same overall $E_{tot}$ and $\mathbf{X}_{tot}$ can be effectively experienced by the system anymore. Thus for instance the energy hypersurface may be partitioned in islands whose boundaries cannot be traversed by the system within times of the order of the laboratory scale. In a structural glass this situation is connected with a diffuse, random, local (in ordinary space) freezing of the configurational degrees of freedom. This aspect will be considered in more detail below, but note that a very important aspect now emerges: the role of time.

### 1.2.2 Partition functions and thermodynamic potentials

The main hypothesis underlying eq. 18 is that each microstate compatible with the given values of the extensive variables is *equally probable* at equilibrium in the whole, isolated system (this has to do with the ergodic hypothesis: the distribution of the representative points of an ensemble at a given time is the same which could be ideally obtained by tracing the positions of the representative point of just one system within a virtually infinite time interval). Besides, the thermodynamic potential in this case is $-S$, because it reaches a minimum at equilibrium.

On the other hand, the extensive variables of any (sub-) system whatsoever undergo continuous fluctuations around the mean, due to the interaction with the environment. So, states with different values of some extensive variable of the (sub-) system do not have in general the same probability, and it is important to find the connection between these probabilities and the "classical" thermodynamic functions characterizing the system (namely the appropriate thermodynamic potential). This is indeed an important point in linear response theory, light scattering, phase transitions, glass transition.

Consider again our whole isolated system as partitioned in a small part ("the system") and the remainder ("the environment"). The total number of microstates $\Omega_{tot}$ can be expressed as the sum of the number of microstates $\Omega_{env}(E_{tot} - E_i;\, \mathbf{X}_{tot} - \mathbf{X}_i)$ of the environment compatible with each *single* microstate "$i$" of the system [characterized by the values $(E_i,\, \mathbf{X}_i)$]:

$$\Omega_{tot}(E_{tot}) = \sum_i \Omega_{env}(E_{tot} - E_i)\,. \tag{19}$$

(the extensive variables $\mathbf{X}$ other than the energy will not be written explicitly from now on for notational convenience; they will be explicitly indicated when needed). For any given state "$i$" of the system, the number of microstates of the environment can be expressed by eq. 18 because a fixed $E_i$ implies a constant (non-fluctuating) $E_{env}$. However, we write the entropy in such a way as to make explicit reference to the *mean* thermodynamic variables $(U,\, \mathbf{X})$ of the system, i.e. $\Omega_{env} = \exp\{k_B^{-1} S_{env}(E_{tot} - U + U - E_i)\}$, and then expand in a Taylor series about these values:

$$\Omega_{env} \simeq e^{S_{env}(E_{tot}-U)/k_B}\, e^{[U-E_i+T\,\boldsymbol{\xi}\cdot(\mathbf{X}-\mathbf{X}_i)]/k_BT} \tag{20}$$

where $\boldsymbol{\xi} \equiv \partial S_{env}/\partial \mathbf{X}$ are the intensive variables of the environment. Since $S_{env}(E_{tot} - U) = S_{tot}(E_{tot}) - S(U)$, substitution back into eq. 19 yields:

$$\Omega_{tot} = e^{S_{tot}/k_B}\, e^{\Psi/k_BT}\, \Xi \tag{21}$$

where

$$\Psi \equiv U - TS - T\,\boldsymbol{\xi}\cdot\mathbf{X} \tag{22}$$

is the thermodynamic potential appropriate for the system and

$$\Xi \equiv \sum_i e^{-(E_i/k_BT + \boldsymbol{\xi}\cdot\mathbf{X}_i/k_B)} \tag{23}$$

is the associated *partition function.*

If no extensive variable is allowed to fluctuate except $U$, then we are in the so called *canonical* representation, and the general

thermodynamic potential coincides with the Helmholtz potential, i.e. $\Psi = U - TS \equiv F$. If instead the only fluctuating extensive variables are the number of particles $N$ and the energy, then $\Psi$ is the Grand potential, and we call this representation *macrocanonical.* Other representations different from these two are possible, and for each one we can readily find the correspondence between the actual potential and the corresponding partition function.

Since $\Omega_{tot} = \exp\{S_{tot}/k_B\}$, it follows immediately from eq. 21 that the thermodynamic potential of a system can be calculated from its partition function as follows:

$$\Psi = -k_B T \ln \Xi \,, \tag{24}$$

that is, the thermodynamic potential is related to a sum over single ("quantum") states among which the system fluctuates.

Last equation can be used to express the entropy of the system in terms of its state probability distribution $p_i$. Indeed,

$$p_i = \frac{\Omega_{env}}{\Omega_{tot}} \simeq \frac{1}{\Xi} e^{-(E_i + T\boldsymbol{\xi}\cdot\mathbf{X}_i)/k_B T} \,, \tag{25}$$

and with the help of eq. 22 we find

$$\sum_i p_i \ln p_i = \frac{1}{k_B T} \left\{ \Psi - \sum_i p_i \left[ E_i + T\boldsymbol{\xi}\cdot\mathbf{X}_i \right] \right\} . \tag{26}$$

On the other hand, the average $\langle \cdot\cdot \rangle$ of any quantity characterizing the state "$i$" of the system can be expressed in terms of the probability distribution by $\langle \cdot\cdot \rangle = \sum_i p_i \, (\cdot\cdot)$. Since $U$ and $\boldsymbol{\xi}\cdot\mathbf{X}$ are the averages of $E_i$ and $\boldsymbol{\xi}\cdot\mathbf{X}_i$ respectively, we find from eq. 22

$$S = -k_B \sum_i p_i \ln p_i \,. \tag{27}$$

This is a very important formula not only from the practical point of view, but also because of its philosophical implications. Indeed, within "classical" thermodynamics, we are used to think of the entropy as a coordinate which, together with the others, contributes

to define completely the thermodynamic state of a system. On the other hand, differently from other cases such as for the energy, for instance, eq. 27 expresses the entropy as an average of a function of the probability itself. Probability, however, is intrinsically associated to a lack of knowledge. This means that in attempting to give a microscopic interpretation of the entropy we are led to the apparent paradox that its microscopic foundation is the lack of information. This contrasts the usual idea aiming to search in the microscopic mechanisms a more firm basis upon which we seek to construct a coherent picture of our macroscopic world. Probably for this reason the entropy leaves that flavor of mystery which almost everyone perceives, the same which can be felt when approaching quantum mechanics.

We now leave aside these philosophical problems to turn back to the calculation of the thermodynamic properties of macroscopic systems. As shown above, the thermodynamic potential can be obtained once the partition function is known. In practical situations one often seeks to find the equilibrium values of some parameter characterizing the system (the distribution of particles among a set of allowed levels, for instance). In this case, a very powerful method consists in the direct maximization of the total entropy of the whole (microcanonical) system. In fact, it is assumed that the environment is quasi-microcanonical, i.e. that the equation $S_{env} = k_B \ln \Omega_{env}$ holds to an extremely high degree of accuracy (in other words this means that $p_i \simeq \Omega_{env}^{-1}$ independent of $i$ is assumed). Then, by direct integration of the equations $T^{-1} = (\partial S_{env}/\partial E_{env})_{\mathbf{X}_{env}}$ and $\boldsymbol{\xi} = (\partial S_{env}/\partial \mathbf{X}_{env})_{E_{env},..}$ we find the relation $\Omega_{env} = K \exp[E_{env} + T\,\boldsymbol{\xi} \cdot \mathbf{X}_{env}]/k_B T$, where $K$ is an integration constant. The total number of microstates accessible to the whole system is thus

$$\Omega_{tot} = \Omega_{env}\Omega = K\, e^{(E_{tot}+\boldsymbol{\xi}\cdot\mathbf{X}_{tot})/k_BT}\, e^{-[U+\boldsymbol{\xi}\cdot\mathbf{X}-k_BT\ln\Omega]/k_BT}\,. \qquad (28)$$

Since $E_{tot}$ and $\mathbf{X}_{tot}$ are fixed, the maximum of $\Omega_{tot}$ can be obtained by minimizing the argument of last exponential in the r.h.s. of eq. 28, i.e. $U + \boldsymbol{\xi} \cdot \mathbf{X} - k_B T \ln \Omega$.

### 1.2.3 Fluctuations

Boltzmann formula carries with it the fact that the extensive variables of a system in contact with an environment fluctuate around mean values which are in fact the equilibrium values of these quantities. So, for instance, the energy of the system changes randomly, but always keeping more or less close to the energy $U$ as defined within classical thermodynamics. For the other extensive variables it is the same. This is a direct consequence of relating the entropy with the number of "possible" microstates, and not with more detailed features of the dynamics characterizing the whole isolated system. We thus considered fluctuations already in eq. 18; this equation was in fact the starting point from which we eventually derived the expression of the thermodynamic potential in terms of the partition function.

The fluctuations of the extensive parameters of the system can be rather easily calculated directly from the knowledge of the partition function. As an example, consider a system whose thermodynamic potential is the Helmholtz free energy. Then, the mean energy fluctuation can be readily found as

$$\langle \delta U^2 \rangle \equiv \langle [E_i - U]^2 \rangle = \frac{\partial^2 \ln Z}{\partial \beta^2} , \tag{29}$$

where $Z = \sum_i \exp^{-\beta E_i}$ and $\beta \equiv 1/k_B T$.

Derivations like this are rather simple, and in general they always are wherever the parameter whose fluctuations we want to calculate does appear explicitly in the expression of the partition function. More generally, the fluctuations of a quantity $w$ whatsoever associated to a (sub-) system can be approximately described by resorting directly to eq. 18. The latter can be inverted to yield $\Omega_{tot} = \exp\{S_{tot}/k_B\}$; then, imposing a well defined value of $w$ to the (sub-) system, thus forcing it to fluctuate within a *limited set of states*, is equivalent to impose an (average) internal constraint to the whole isolated system. This reduces the number of microstates available, and the total entropy too. Since the probability $p(w)$ of

observing the value $w$ is proportional to the number of accessible microstates compatible with this average internal constraint, then the deviation from the mean $\langle w \rangle$ can be related to the corresponding changes induced on $S_{tot}$. The problem is thus reduced to find the functional dependence of $S_{tot}$ on the actual value $w$ assumed by the variable of the system.

As an example, let $G$ be the average Gibbs potential of a system, and let $S_{tot}(G)$ be the corresponding overall equilibrium entropy. Now, among the whole number $\Omega_{tot}$ of states at equilibrium, there will be a fraction of them for which the Gibbs potential of the system takes a new value $G'$. Their number is $\Omega'_{tot} = \exp\{S_{tot}(G')/k_B\}$, and the probability for this new value to occur as a result of a fluctuation is

$$p\,(\Delta G) \propto \frac{\Omega'_{tot}}{\Omega_{tot}} = e^{\,\Delta S_{tot}/k_B}\,, \tag{30}$$

where $\Delta S_{tot} \equiv S_{tot}(G') - S_{tot}(G)$ and $\Delta G \equiv G' - G$. Now, assuming that no work other than $p\,dV$ is done by the system during the fluctuation, from eq. 13 we find

$$p\,(\Delta G) \propto e^{\,-\Delta G/k_B T}\,. \tag{31}$$

In the general case, we may use the following approximate procedure. The entropy of the whole isolated system, $S_{tot} \equiv S_{tot}(\langle w \rangle)$ is indeed a maximum at equilibrium. Since the system is a much smaller part of the whole, then a truncated Taylor expansion suffices to describe the effect of the deviation $w - \langle w \rangle$, and we may take

$$S_{tot}(w) \simeq S_{tot}(\langle w \rangle) + \frac{1}{2}\frac{d^2 S_{tot}}{d\,w^2}(w - \langle w \rangle)^2 \tag{32}$$

(gaussian approximation). The average square fluctuation amplitudes can then be readily obtained from the expression

$$p(w) \propto e^{-|d^2 S_{tot}/d\,w^2|(w-\langle w \rangle)^2/2k_B}\,, \tag{33}$$

i.e. explicitly,

$$\langle \delta w^2 \rangle = k_B \left| \frac{d^2 S_{tot}}{d\,w^2} \right|^{-1}\,. \tag{34}$$

# 2 Polymers

## 2.1 A brief introduction

Polymers are very special (macro-) molecules, characterized by a huge number of internal conformational degrees of freedom. One polymer molecule can thus be viewed as a macroscopic system to which thermodynamics can be applied. Figure 2 shows part of a polyethylene chain; the degrees of freedom are the rotations about each C—C covalent bond. About calling these degrees of freedom "internal", however, things are not as simple as it may appear. The shape of a macromolecule, depending on the actual values of the set of its rotational coordinates, is indeed central with respect to entanglements, which represent a relevant contribution to the overall inter-molecular interaction mechanisms in polymer melts. In

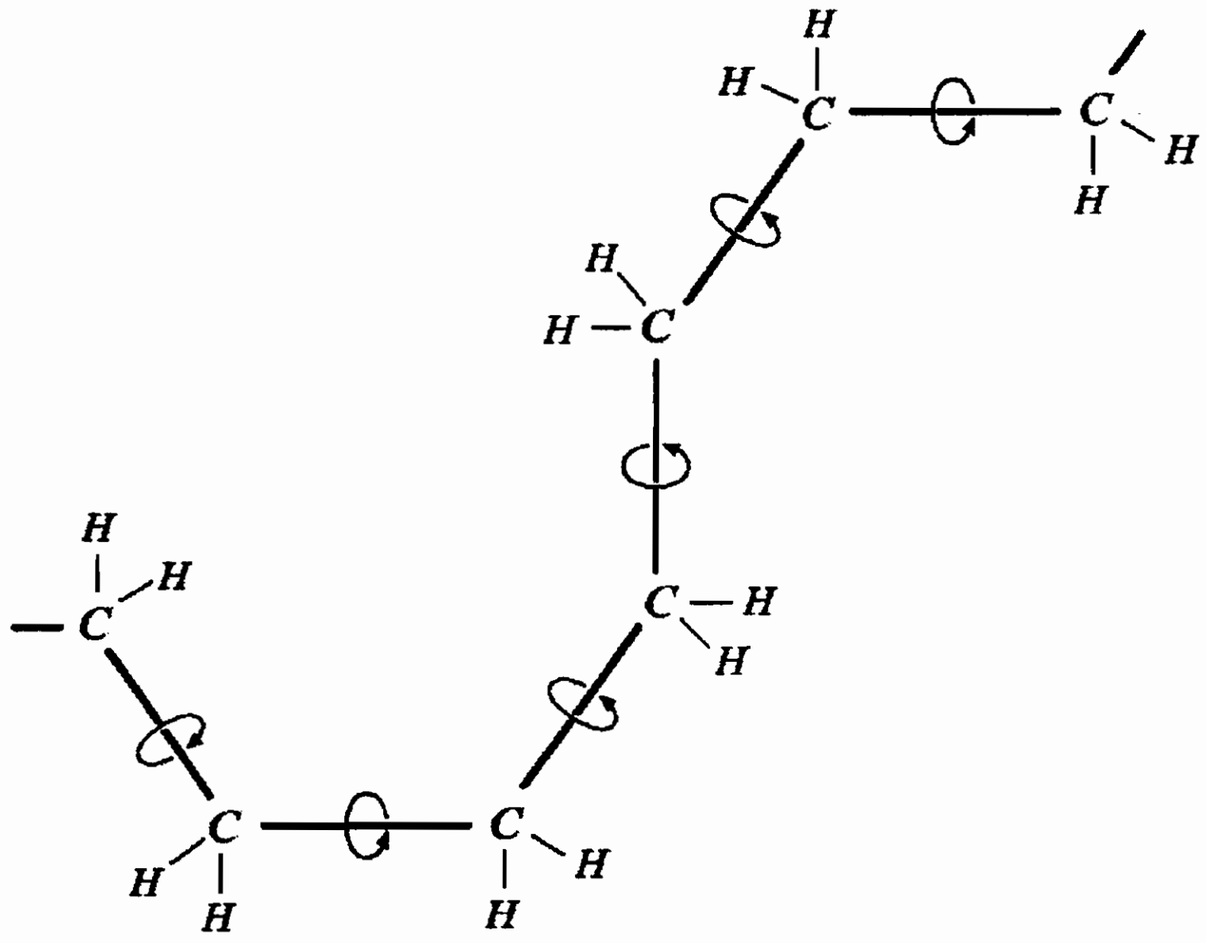


Figure 2: The overall chain conformation is a function of the values of the bond rotational degrees of freedom in a single polyethylene chain

this respect the degrees of freedom of a single macromolecule cannot be considered internal, and are means through which molecular conformations mutually interfere.

This first introduction hints at the possibility that, with polymers, things may be complicated (as we shall see indeed). These systems are characterized by a very rich phenomenology (see for instance refs. [2, 3], and we shall limit ourselves to consider some particular aspects of their complex behavior, namely, within the context of their (pseudo -) phase transitions.

We start from a thermogram typical of polyethylene terephtalate (PET), a widespread used crystallisable polymer (plastic bottles for water are made of PET). By rapid temperature quenching from above the melting point, this material can be rather easily driven to a completely glassy state. The thermogram shown in Fig. 3 has been obtained by re-heating PET from this glassy state. The ordinates of this plot represent the heat flowing into the sample $dH/dt$. This quantity is in fact proportional to the specific heat at constant pressure $C_p$ because

$$C_p = \frac{dH}{dT} = \frac{dH}{dt} \cdot \frac{dt}{dT} \tag{35}$$

and the heating rate $dT/dt$ is a constant. On heating the sample, a first endothermal process is detected in the temperature range $70 \lesssim T \lesssim 85$ °C. The specific heat in the high-$T$ side of the process is significantly upward shifted with respect to its value at the other side. In this temperature region, the system passes from a quenched, frozen-in disorder typical of a glass, to a state where still no molecular order sets in, but fluctuations start to be active (thus causing an increase in the specific heat). If in these conditions the system were cooled again, then the fluctuations would freeze-in again. This temperature interval is referred to as the *glass transition* region.

Upon heating the system further, we can observe a very sharp exothermal process starting from about $T = 140$ °C, which is ascribed to crystallization. It is worth to be mentioned that the peak

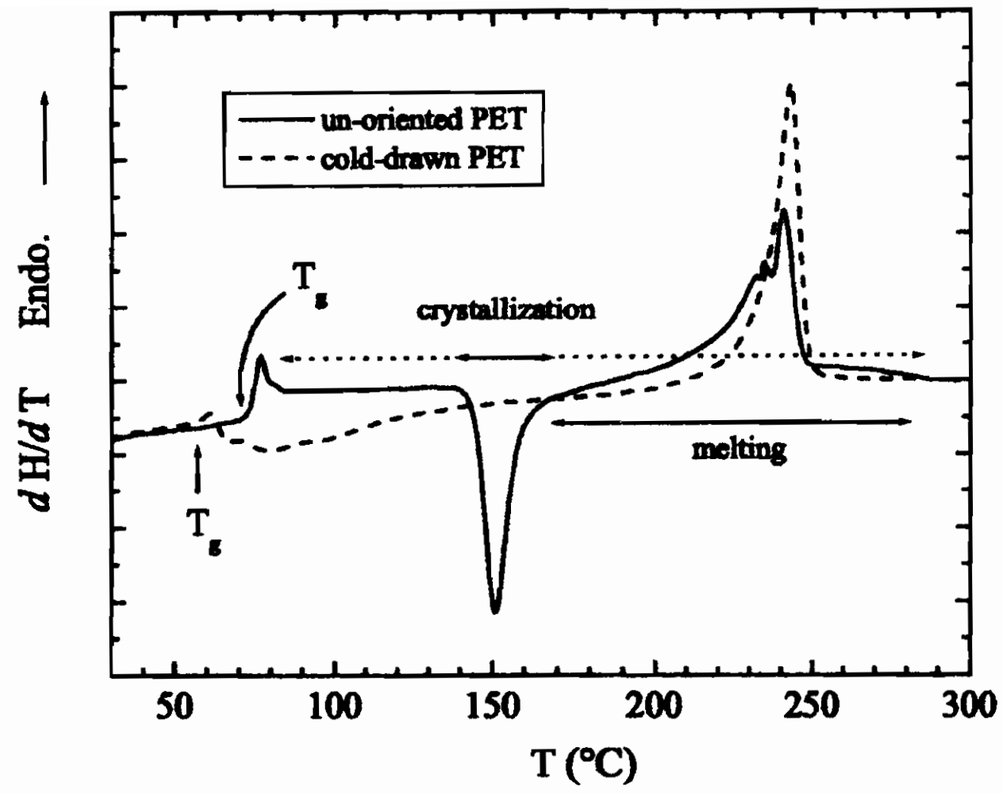



Figure 3: Heat flux as a function of the temperature during the heating of an initially glassy un-oriented PET sample (heating rate $dT/dt = 40\,°\mathrm{C/min}$. First the glass transition process takes place ($T_g \sim 75\,°\mathrm{C}$), followed by a crystallization exotherm ($T \sim 120\,°\mathrm{C}$). Upon further heating, a rather wide endotherm, ending at about $T \sim 285\,°\mathrm{C}$ reveals the melting of unstable crystals.

location does not have the same implications as in low molecular weight substances. If one waits long enough, crystallization can be observed also at lower temperatures (above the glass transition, of course).

Beyond the high temperature side of the crystallization peak, we find a rather wide endothermal peak. In this region unstable crystallites melt and re-crystallize, giving rise to a very complex behavior. The overall process, however, is the melting of all crystalline domains, which completely disappear at a temperature $T_m^{\infty} \approx 285\,°\mathrm{C}$.

Figure 4 reports the results of a Dynamical Mechanical Thermal Analysis (DMTA) carried out on the same system, with the same thermal history. With this technique it is possible to observe some internal microscopic processes through the macroscopic me-

chanical response to a force applied on the sample, as a function of the frequency. For completeness, a very brief account of the basic principles of this technique is given (see for instance ref. [2] for a more extended treatment.

A sample acted upon by a force $f$ will respond by changing its shape by an amount $s$. Assume that in quasi-static conditions (i.e. relative to the microscopic relaxation times characteristic of the sample) these two quantities are linearly related to one another by $s = K_{st} f$, and take this condition as a reference state for some actual non-zero force $f_0$. Then, if the latter is suddenly varied by an amount $\delta f$, the shape $s$ will also change to adjust to a new equilibrium state. If $\delta f$ is small enough, the transient variation of $s$ can be well approximated by the equation

$$\frac{d\,s}{d\,t} = -\frac{1}{\tau}[s - K_{st}(f_0 + \delta f)]\,, \tag{36}$$

where $\tau$ is the characteristic time over which the material changes its shape, and is related to the dynamics of the (relevant) microscopic modes within the material itself. If the system is in thermodynamic equilibrium, this result can be proved by the fluctuation-dissipation theorem, but we must skip this aspect.

If $\delta f \sim \exp\{i\omega t\}$, then $\delta s = K_{st}[1 - i\omega\tau]^{-1}\delta f$. The *dynamic compliance*

$$J \equiv \frac{\delta s}{\delta f} \tag{37}$$

then carries information on the molecular dynamic modes of the system. The real and imaginary parts of the compliance give rise to an *in-phase* and an *out-of-phase* responses of the system. The latter is connected to the amount of mechanical work converted into heat by the system. Indeed, the average dissipated power can be found to be related to $J$ by

$$\langle \dot{W} \rangle \propto \Im(J)\,. \tag{38}$$

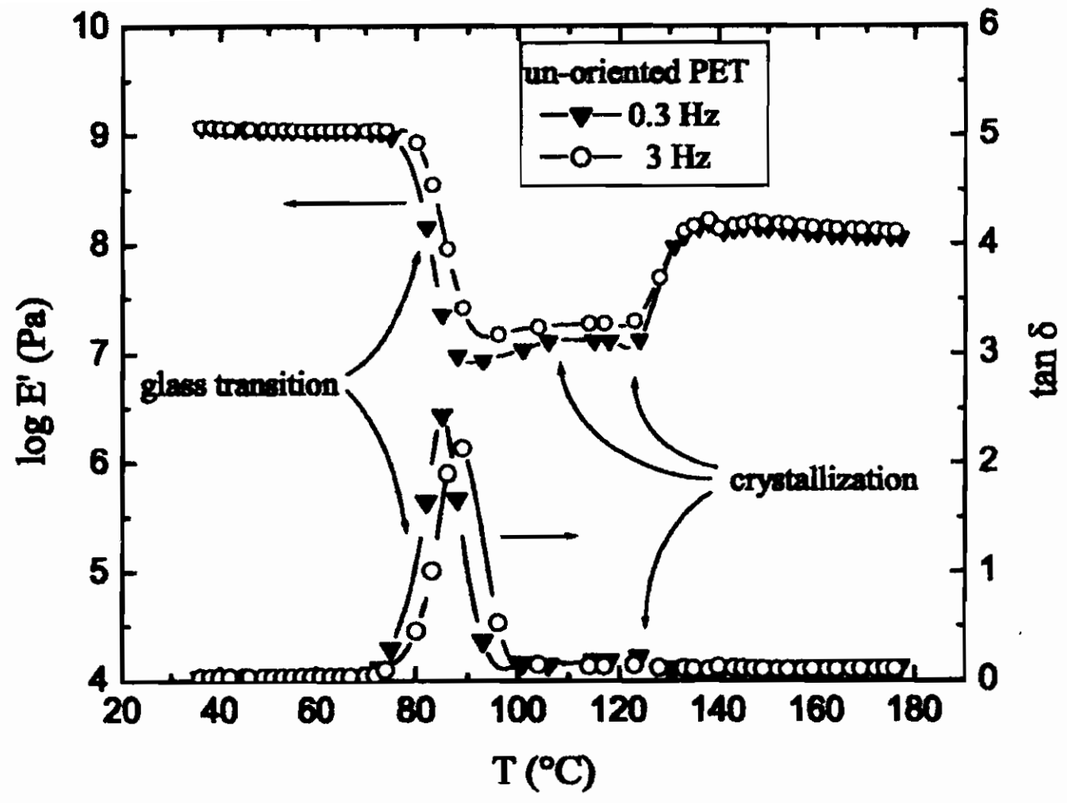



Figure 4: Storage modulus and $\tan\delta$ for un-oriented PET at different frequencies.

A very useful quantity for the analysis of the system is the ratio

$$\tan\delta \equiv \frac{\Im(J)}{\Re(J)}\,. \tag{39}$$

The glass transition region is revealed by a rather wide and pronounced peak when $\tan\delta$ is plotted against the excitation frequency. The actual frequency interval where this so called $\alpha$-process is observable, depends on the temperature of the sample. In practice, since the relaxation processes giving rise to the $\alpha$ dispersion have characteristic times which depend on the temperature, DMTA measurements are carried out with temperature scans at fixed frequencies, and the $\alpha$ dispersion is revealed by a wide $\tan\delta$ vs. $T$ peak, with a position depending on the frequency of the mechanical excitation. This dependence is the fingerprint of the glass transition process and is connected with its kinetic character. Remind that we gave a little hint about the increasingly important role of time when describing the motion of the representative point of the system in a partitioned phase space (see last comment before section 1.2.1). By

varying the frequency of the excitation, we change the characteristic laboratory time with which the internal relaxational processes of the system compares; this is the physical reason of the dependence of the peak location on the frequency. In the region where the loss peak appears, the rigidity of the material drops significantly, as can be seen from the storage (in phase) modulus reported in Fig. 4.

Upon further heating, crystallization starts. This causes the storage modulus to increase sharply again. The step location does not in this case depend on the frequency; this is indicative of a true thermodynamic phase transition.

In this chapter we shall be concerned with some aspects of the glass transition and the crystallization processes.

## 2.2 A short account of the Structural Glass Transition, the Adam-Gibbs model

When introducing the effect of the internal constraints on the overall entropy in an isolated system, we digressed about the possibility that some regions of the phase space may become inaccessible to the representative point of the system, even if compatible with the overall values of the extensive variables. This possibility is not just academic, and actually represents a point of view by which the widespread and yet not well understood process of the structural glass transition is currently investigated. It is not the point here to give a thorough picture of the complexity concerning the glass transition; rather we shall define a small number of firm concepts in the light of which some interesting experimental results can be discussed.

Vitrification is a process by which a liquid transforms into a solid without showing afterwards any structural regularity at the atomic level, as would happen in the formation of an ordinary crystal. As shown in Fig. 5, the formation of a glass by cooling is monitored by a more or less rapid decrease of the specific heat $C_p$ within a comparatively small interval around a temperature $T_g$ (the so called *dynamic* glass transition temperature). Grossly speaking,

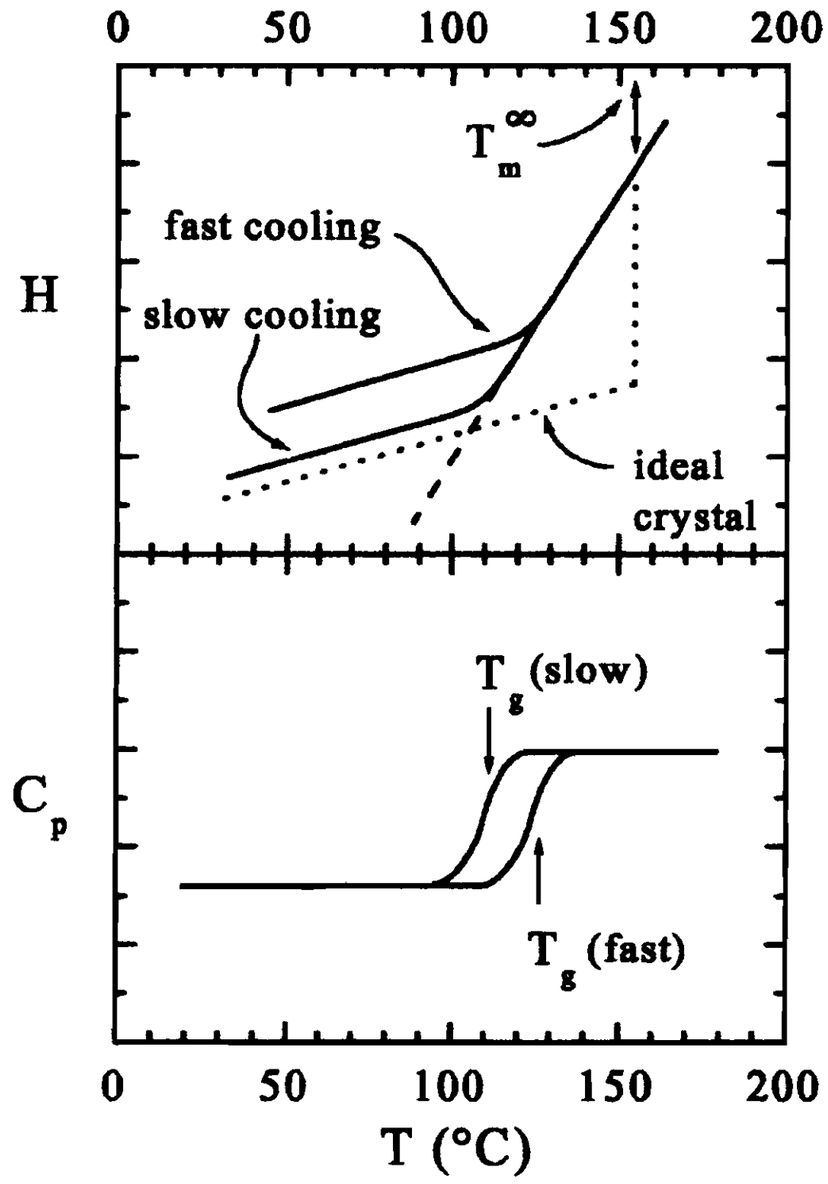



Figure 5: Changes of enthalpy, $H$, and specific heat at constant pressure, $C_p$, at the glass transition for fast and slow coolings of a liquid. The *dynamic* glass transition temperature, $T_g$, is a function of the cooling rate.

the drop in $C_p$ can be associated to the freezing-in of the large scale (configurational) motions such as density fluctuations; instead, the vibrational degrees of freedom contributing to $C_p$ as in an ordinary crystal remain active. For this reason, the overall system is approximately viewed as consisting of two intermingled subsystems: one is associated to the vibrational degrees of freedom, the other to the configurational ones.The total partition function is assumed to be roughly given by the product of those of the vibrational and the configurational subsystems.

Concerning the enthalpy, the final values down to which $H$ drops, are close to but slightly larger than those typical of the

crystal phase of the same material at the same temperature (remind that a system can always undergo a glass transition, even in the case that its crystalline state does not exists, such as some polymers). One problem is that for any given substance the temperature region where this process takes place, as well as the $H$ values in the final glassy state, depend on how fast the liquid is cooled. Larger cooling rates drive the transition at higher temperatures, with higher residual values of $H$ below $T_g$ (compare with Fig. 4, where the $\alpha$-peak shifts to higher temperatures as the frequency increases). These experimental facts suffice for us to point out a significant difference from ordinary phase transitions, namely, that *Time* now enters a class of phenomena whose description was the domain of thermodynamics: a language where infinite lapses of time (e.g. in reversible processes) were the rule.

When *observing* a glass transition process, we must always keep in mind that a confrontation between characteristic laboratory time scales and natural speed of the process itself is underlying. This is to point out the *kinetic* character of the observation. The following problem now emerges naturally: If the trend of the enthalpy of a supercooled (but crystallisable) liquid were to be extrapolated towards the absolute zero (dashed line in Fig. 5), the entropy would drop below that of a crystal and eventually become even negative *if* the glass transition would not intervene. So, a kinetic process would prevent the breakdown of thermodynamics ! This is known as the Kauzman paradox, and is indeed a good reason to investigate about the possibility that a true thermodynamic transition exists at a temperature $T_K$ (i.e. the *ideal* glass transition temperature), which is hidden by the kinetic interference of the (dynamic) glass transition observed about the higher $T_g$. In fact the existence of such a transition has been demonstrated for polymers [4], but not in general for other low molecular weight substances.

Once this brief introduction has been given, we must now leave these fundamental problems to consider, more pragmatically, the observation of the glass transition within the limits which our embedding world has established for us.

Structural relaxation is a process whereby the (local) configuration of a group of molecules changes. Its characteristic time scale thus influences directly the viscosity $\eta$ of a liquid, being the latter a measure of the resistance to flow. In fact, the reduction of the specific heat on approaching the *dynamic* glass transition at $T_g$ is always accompanied by a huge increase of the viscosity. In a number of systems, for instance, the way the viscosity increases with decreasing $T$ fits the law $\eta \sim \exp\{A/(T-T_K)\}$. The measurement of $\eta$ cannot be performed down to temperatures even close to $T_K$ because the dynamic glass transition at $T_g > T_K$ intervenes. The temperature $T_K$ could in this case be associated to an *ideal* glass transition temperature in the spirit of Kauzman arguments.

In the following we shall outline the theory developed by Adam and Gibbs [5] to describe the divergence of $\eta$. It represents a particularly interesting and fundamental attempt to describe in general the approach to $T_g$, from the liquid state, in terms of known thermodynamic concepts. Apart of this, its relevance in the present exposition is in that it introduces some ideas which will be useful for us in developing a theory for the vitrification process observed in confined amorphous PET.

Central to the theory is the concept of Cooperatively Rearranging Region (CRR), that is, a group of molecules (or monomers) able to rearrange into a different configuration *independently of its environment* as a result of an energy and/or volume fluctuation. Cooperativity means that the motion of a molecule whatsoever is functional to the motion of the others, to circumvent steric hindrances and allow for a change of the overall configuration. We consider the (sub-) system as consisting of $M$ distinguishable CRR's (all made of $n$ molecules), for any of which the partition function associated to the Gibbs potential (i.e. the potential associated to the possibility that both the volume and the energy of the system fluctuate) is

$$Z(n,p,T) = \sum_{E,V} \omega(n,E,V)\, e^{-(E+pV)/k_B T}\,, \tag{40}$$

where $\omega(n, E, V)$ is the degeneracy of the CRR state with given $E$ and $V$. Among the states accounted for in the calculation of $Z$, consider the set of the $E$ and $V$ values allowing for a structural rearrangement (i.e. those for which the configurational component of $\omega$ is larger than one) and construct the associated *reduced* partition function $Z'$ by limiting the summation to just these values. Then, the fraction $m/M$ of CRR's which are able to undergo an actual configurational change is:

$$\frac{m}{M} = \frac{Z'}{Z} = e^{-(G'-G)/k_B T} . \tag{41}$$

The cooperative transition probability $p(T, n)$ is proportional to this number. Let $A$ be this proportionality factor; then, accounting for all possible sizes of the CRR's, we find for the overall transition probability

$$p\,(T) = \sum_{n \geq n^*} A\, e^{-n\delta\mu/k_B T} \approx \overline{A}\, e^{-n^*\delta\mu/k_B T} , \tag{42}$$

where $n^*$ is the minimum number of molecules for a group to be a CRR, $\delta\mu = (G' - G)/n$ is the chemical potential change associated to each molecule to overcome the energy barrier hindering rearrangement, and we have assumed $\delta\mu \gg k_B T$ to approximate the summation. Equation 42 implies that the overwhelming majority of transitions is performed by the smallest CRR's. Note that $\delta\mu > 0$; this is consistent with the general fact that when some constraint to the configurational motion (i.e. to the accessible microstates) *appears* or becomes effective in correspondence with the $E$ and $V$ values which *do not* allow rearrangement, then the appropriate thermodynamic potential increases (in this respect, remind Section 1.1.2 above and eqs. 18 and 24).

In eq. 40 all degrees of freedom are accounted for, those describing the vibrational motion of molecules around their local potential energy minima, as well as those associated to the exploration of the "configurational islands" in which the whole phase space is partitioned. In an ideal crystal the latter coalesce in just one point.

The basic idea is that in supercooled liquids the overall partition function can be (approximately) considered to be the product of a vibrational contribution, the same we would have in a crystal, and a configurational one. In other words, the liquid is viewed as a composite system consisting of one part collecting the totality of the vibrational microstates, and another one whose only degrees of freedom are configurational. This idea is intrinsic to the hypothesis that the drop of $C_p$ at the glass transition is a consequence of the structural arrest, and that the main contribution to $C_p$ in a glass comes from the same mechanisms as in a crystal (see above).

In the light of all this, the overall *configurational* entropy $S_c$ results from the sum of the configurational entropies $s_c^*$ of all minimum sized CRR's (which we remind are distinguishable and independent from one another), and considering an Avogadro number $N_A$ of molecules we find

$$n^* \simeq \frac{N_A s_c^*}{S_c} \tag{43}$$

[as an estimate for an order of magnitude, $s_c^* \sim \mathcal{O}(k_B \ln 2)$ since the minimum number of configurations a CRR may have is two]. Note that $n^*$ diverges as $S_c \to 0$; this is in fact the mechanism to which the divergence of the viscosity is ascribed when the system cools down ($\delta\mu$ remaining fixed). Since the viscosity of a liquid is proportional to the characteristic structural relaxation time, and the latter is proportional to the inverse average probability of the cooperative rearrangement, eq. 42, it turns out that the viscosity is related to the configurational entropy by

$$\eta \propto p^{-1}(T) \propto \exp\left\{\frac{B}{TS_c}\right\} . \tag{44}$$

The actual temperature dependence of the viscosity depends on the properties of the liquid which is being considered, i.e. the actual form of $S_c(T)$; this aspect is of no concern for us presently. The main significance of last equation is the relationship between a thermodynamic quantity ($S_c$) and the relaxation kinetics of a system (e.g. $\eta$), which is ultimately observed in real experiments.

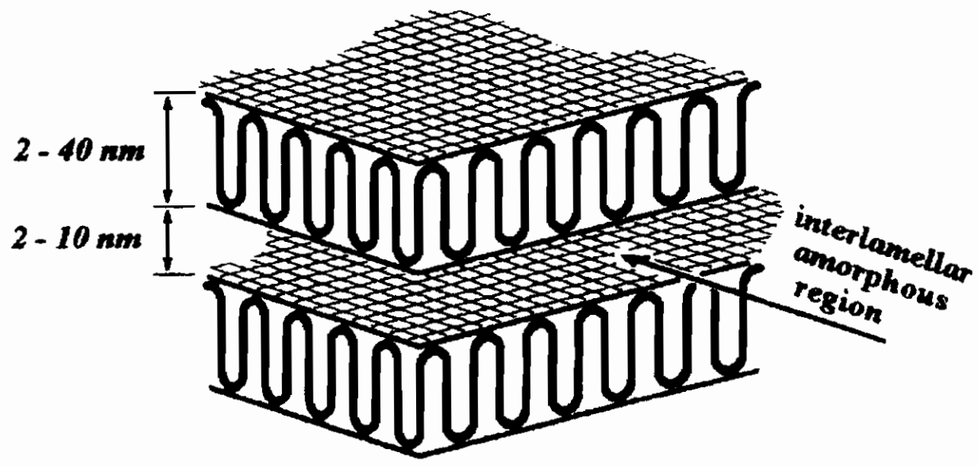
We note that eq. 44 above (and its derivation) does not need that the glass transition is approached by cooling the system. Other mechanisms, such as chemical vitrification (polymerization) or the confinement of an amorphous polymer into small regions, can be responsible of the progressive decrease of configurational entropy, also at a constant temperature. The latter mechanism will be explicitly considered in the sequel; however, since the confinement mechanism to be considered in detail is strictly connected with the presence of a crystalline phase, we must briefly outline the crystallization process taking place in polymeric systems.

## 2.3 Crystallization of Polymers

Crystallization may occur in polymers; it is a very complex process, and the details of the mechanisms involved are still matter of debate. Here we shall only give a rather naïve picture of it, moreover focusing only on those aspects which will be useful for us.

Due to the (long) chain-like structure of the molecules, entanglement interactions affect crystallization considerably. The long times required for chain disentanglement render the extended chain crystallization impossible, unless the chains are short enough. For the same reason, a polymer cannot crystallize completely. Kinetics however plays a role, and finite time crystallization results in the formation of stacked lamellar crystallites separated by amorphous layers, as shown by Fig. 6. It is evident that the surface tension associated to so many crystal/amorphous interfaces, increases the overall final free energy significantly above the value which could be reached if the system were to crystallize in an extended chain mode. This circumstance is reminiscent of the structural glass transition considered above: reaching the lowest free energy state would require an infinite time; paths towards different (less favorable) *thermodynamic* states, which can be reached within finite times, are thus naturally preferred (are in fact what we are let to see) within our time scales. Note that, like in glass formation, also in this case the problem is connected with configurations.

a)




b)

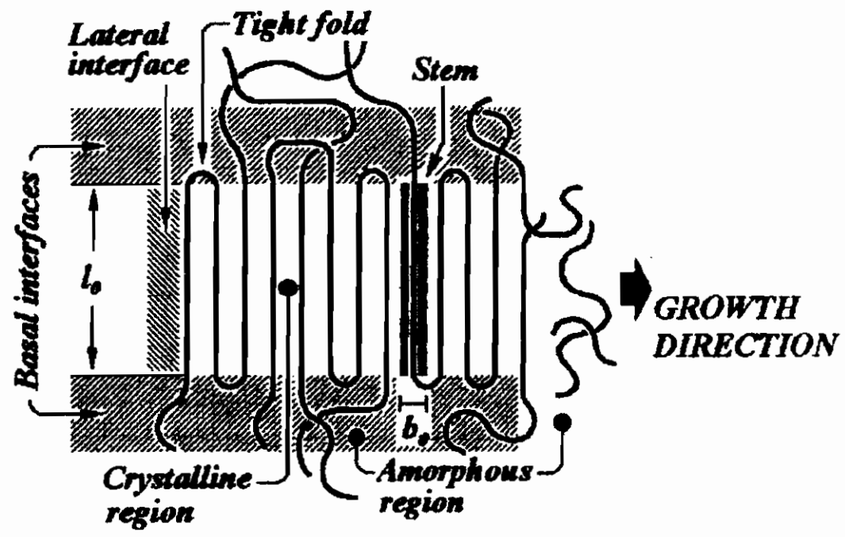



Figure 6: **a**) Crystalline lamellae arrange in stacks on an intermediate length scale (schematic); the interlamellar region is amorphous, it contains a number of defects which prevent further crystallization (thickening). To make the drawing simpler, all chains emerging from the basal planes tight-fold to re-enter nearby in the same lamella. Actually, things are more complicated: a number of chains leaving a basal plane can either re-enter the same lamella (but not necessarily with a tight fold), or enter into another one (tie chains). **b**) Side view of a lamella undergoing lateral growth.

In general, a more or less deep undercooling must be necessarily imposed to a polymer melt in order to drive crystallization within finite times (provided the temperature is not too close to $T_g$). We consider now the crystallization mechanism where phase separation starts by *nucleation* from the metastable liquid [2, 6]. Free energy *fluctuations* may induce the formation of crystalline nuclei at a rate

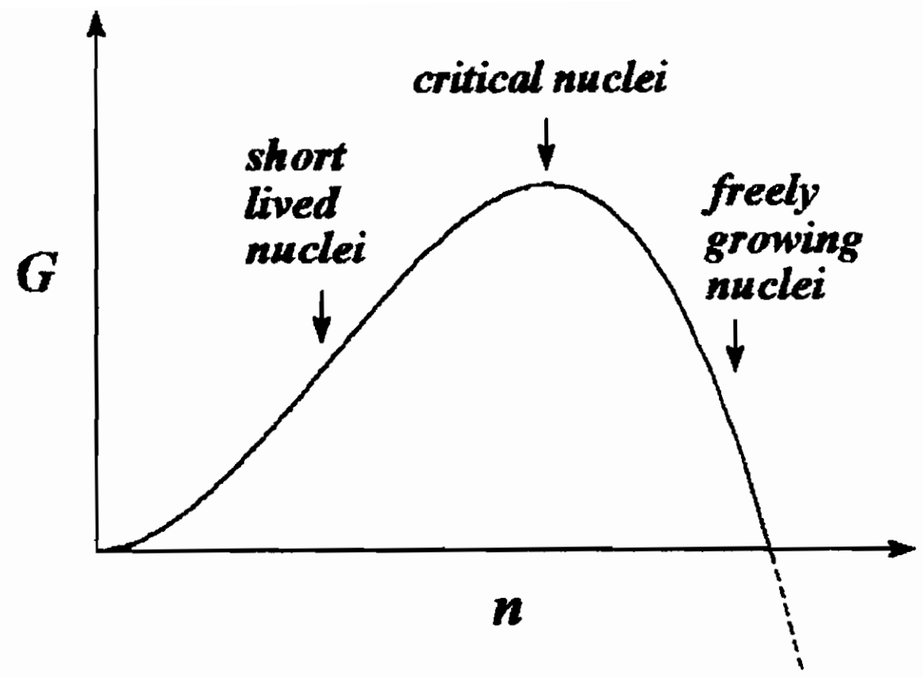



Figure 7: Free energy $G$ associated to the formation of a crystalline nucleus. The functional dependence on the number of units $n$ forming the nucleus, is as follows: $G = -a\,n + b\,n^{2/3}$, with $a$ and $b$ positive constants.

proportional to the probability

$$p \sim e^{-G/k_B T}\,, \tag{45}$$

where $G$ is the Gibbs free energy associated to the formation of a nucleus. Since the latter is a small system, surface effects are important. So, in addition to a term proportional to the number of molecules of which the nucleus is made, $G$ must also include a (contrasting) contribution proportional to the overall energy needed to form the interface area $s$:

$$G = n\Delta\mu + s\sigma\,, \tag{46}$$

with $\Delta\mu \equiv \mu_c - \mu_a < 0$ the chemical potential difference between the crystalline and the amorphous phases, and $\sigma$ the surface tension. If the nuclei are large, then the further increase of their dimensions is thermodynamically advantageous, if they are too small then they will tend to disappear (see Fig. 7).

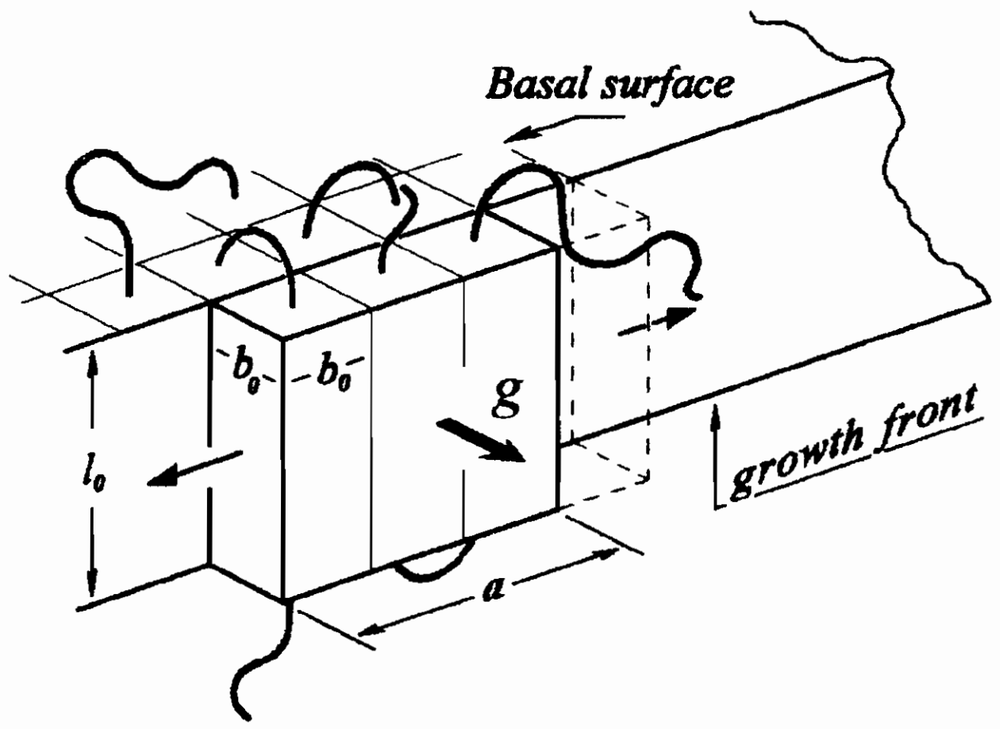



Figure 8: A secondary, slightly supercritical nucleus at the growth front of a lamella. Chain tight folds and/or entanglements (not shown) in the amorphous phase adjoining the basal planes contribute to hinder growth along a direction perpendicular to the basal planes themselves, so the secondary nuclei grow along the directions indicated by the small arrows eventually forming a new, extended layer. This latter process is much faster than secondary nucleation, and the overall increase of the lamellar extension (locally indicated with an arrow labelled by a **g**) takes place with a speed controlled by the secondary nucleation rate.

We shall now focus on the growth process of the crystalline lamellae, so to relate their morphology to the relevant thermodynamic parameters. There are two main growth modes, namely, *lateral growth*, which will be introduced just below, and *lamellar thickening*, which will be our concern afterwards, when considering the glass transition induced by confinement in the interlamellar regions of an oriented polymer.

Lateral growth consists in the increase of the basal interface extension of the lamellae (see Fig. 6), their thickness remaining substantially unaltered. At moderate undercoolings this process is governed by *secondary nucleation*, consisting in that a nucleus

forms (attaches) at some point of the lamellar front; then, if it is large enough, it will grow very fast along the growth front of the lamella as indicated schematically in Fig. 8. The Gibbs free energy of a secondary nucleus consists again of a bulk term involving the chemical potential change $\Delta\mu$, together with two *basal* and two (not four) *lateral* interfacial contributions, i.e.

$$G = v_c^{-1} V \, \Delta\mu + 2 b_0 a \, \sigma_e + 2 b_0 l \, \sigma \, , \tag{47}$$

where $v_c$ is the molar volume in the crystalline phase, $b_0$ is the stem diameter, $a$ is the lateral dimension of the nucleus, $l$ is the stem length and $V = a \, b_0 \, l$; $\sigma$ and $\sigma_e$ are the lateral and basal specific surface free energies (see Fig. 6). Critical nuclei (see Fig. 7) are those for which both the following conditions hold:

$$\frac{\partial G}{\partial a} = 0 \tag{48}$$

$$\frac{\partial G}{\partial l} = 0 \, ; \tag{49}$$

the Gibbs free energy associated to their formation is thus

$$G_0 = \frac{4 v_c b_0 \sigma \sigma_e}{|\Delta\mu|} \, . \tag{50}$$

Nuclei which are slightly larger than critical will grow (laterally) very fast compared to the secondary nucleation rate (that is, secondary nucleation is rate determining process for the overall increase of the lamellar extension; but of course this is true only for moderate undercoolings). Growth along the direction perpendicular to the basal planes, however, is significantly hindered by the entanglements in the adjoining amorphous regions, that is why the supercritical nuclei will substantially tend to increase in dimension laterally. This means that the final lamellar thickness will remain very close to the value obtained from eq. 48:

$$l_0 \simeq \frac{2 v_c \sigma_e}{|\Delta\mu|} \, . \tag{51}$$

Since the difference between the chemical potentials in the crystalline and in the amorphous phases depends on the undercooling, this equation relates the thickness of a lamella to the thermodynamic conditions under which it has grown. Note that at the equilibrium melting point, $T_m^\infty$, the chemical potential difference vanishes and the lamellae would be infinitely thick (thus requiring an infinite time to form).

The above expression can be made more clear and useful as follows. At equilibrium, coexistence of the crystalline and amorphous phases is described by the condition $\Delta\mu = 0$ i.e. by $\Delta H_f = T_m^\infty \Delta S_f$, where $\Delta H_f$ and $\Delta S_f$ are the molar melting enthalpy and entropy at equilibrium (remind eq. 14). Then, if crystallization is driven at a temperature $T < T_m^\infty$, the chemical potential difference can be approximately obtained by direct integration of the Gibbs-Duhem relation (remind eq. 10) $d\,\Delta\mu = -\Delta S_f\, dT$. Assuming $\Delta S_f \approx$ const. we find

$$\Delta\mu \simeq \Delta H_f \frac{T - T_m^\infty}{T_m^\infty} \,. \tag{52}$$

This relation can be readily inserted into eq. 51 to yield the well-known Gibbs-Thomson equation

$$T \simeq T_m^\infty \left[1 - \frac{2v_c\sigma_e}{l_0 \Delta H_f}\right] \,, \tag{53}$$

establishing an approximately linear relationship between the temperature $T$ at which crystallization is driven, and the inverse (average) thickness of the lamellae. A test for this equation requires the use of non-trivial techniques, such as Small Angle X-ray Scattering (SAXS). In many investigations this relation is trusted at the outset (for instance because previous checks were performed elsewhere on the same material), and used to infer $l_0$ or other quantities such as the equilibrium melting temperature $T_m^\infty$. We shall see how the Gibbs-Thomson equation will be important in the forthcoming.

Equation 51 is a consequence of a balance condition. If the lamellae are growing at a given undercooling, it is possible to re-melt them upon a slight increase of the temperature. In this case,

recrystallization into (slightly) thicker structures may take place. The opposite, however, does not occur. If a crystalline lamella has grown at a given undercooling, then further decrease of the temperature only renders it thermodynamically more stable, and no thinning process takes place.

Still concerning eq. 51, we have to keep in mind that in principle lamellar thickening would be favored by the corresponding decrease of the overall free energy of the system. As mentioned above however, this process is largely hindered by entanglements trapped into the interlamellar regions; moreover, chains leaving the basal surface often re-enter nearby into the same lamella, thus forming *tight folds*. The only possible process occurring in the basal interphase is thus a slow reduction of the surface roughness.

## 2.4 A digression about the basal surface tension

Both the lamellar thickness $l_0$ (eq. 51) and the crystal growth rate

$$v \propto p\,(G_0) \sim \exp\left\{-\frac{4v_c b_0 \sigma \sigma_e T_m^\infty}{k_B T |\Delta\mu|}\right\} \tag{54}$$

depend linearly on $\sigma_e$. So, if we are able to change $\sigma_e$ somehow, we should observe consistent variations in $v$ and $l_0$ accordingly.

The value of $\sigma_e$ can be varied by introducing small amounts of a nucleating agent dispersed in the melt. The changes induced in $v$ and $l_0$ can be monitored by calorimetry [7] and by SAXS respectively.

What should we expect? It is known that during crystal growth the molecules of a non co-crystallizing nucleant leave the amorphous regions (adjoining the lamellar front) which are about to crystallize. This is shown schematically in Fig. 9. These nucleant molecules remain close to the basal surfaces because they cause a slight local decrease of $\sigma_e$ (see below). Their final (basal) surface density is directly proportional to the crystallizing volume, i.e. to the lamellar thickness because the stem diameter is a constant.

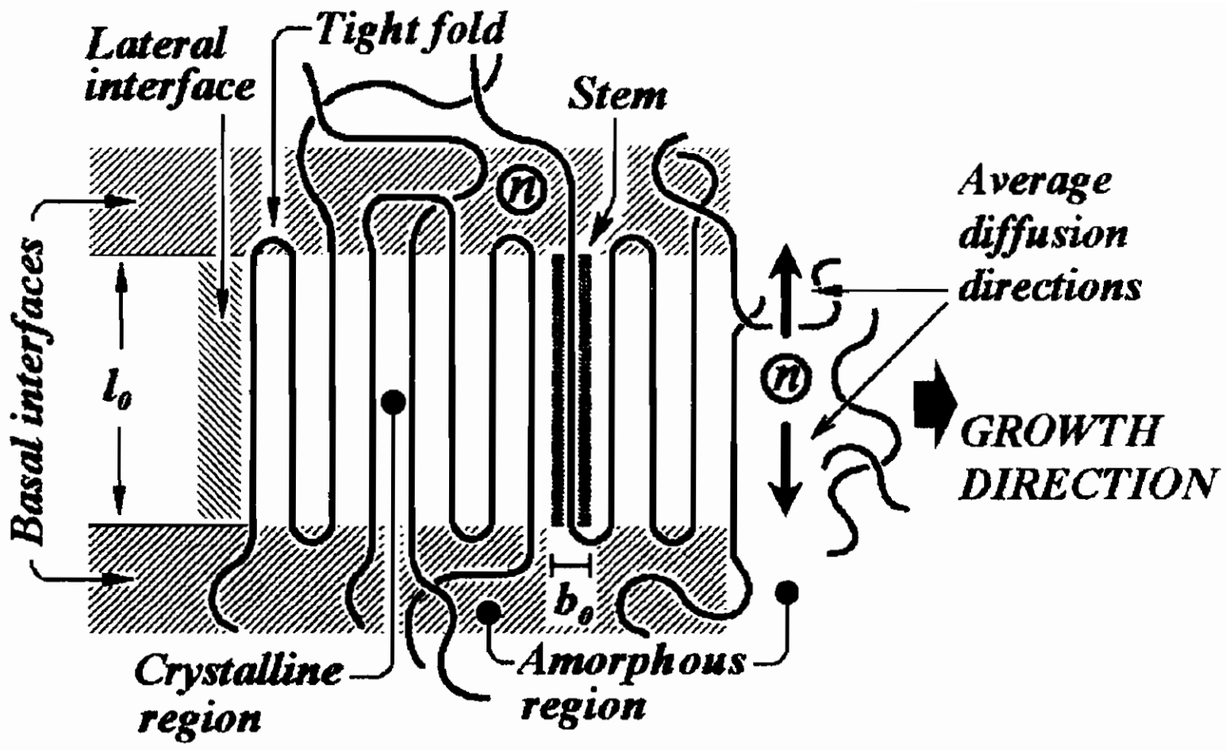



Figure 9: The molecules of a nucleating agent leaving the amorphous region adjoining the growth front during lateral growth

If the concentration of the nucleant molecules is small enough, we can approximately calculate their final surface density taking $l_0$ in place of the actual lamellar thickness. Since $l_0$ is inversely proportional to the undercooling $\Delta T$ (see eqs. 51 and 52), the actual basal surface tension $\sigma_e$ in the presence of a nucleating agent is dependent on the temperature:

$$\sigma_e \simeq \sigma_{e0} \left(1 - \frac{a}{\Delta T}\right) , \tag{55}$$

where $\sigma_{e0}$ is the basal surface tension in the absence of the nucleant, and $a$ is a phenomenological coefficient.

If in place of a constant $\sigma_e$ we substitute eq. 55 into eq. 51, we find

$$l_0 \simeq \frac{2 v_c \sigma_{e0}}{|\Delta \mu|} \left(1 - \frac{a}{\Delta T}\right) . \tag{56}$$

Figure 10 reports the data fitting of the linear growth velocity of the lamellae in polyethylene oxide (PEO) nucleated with a number density of $2.6 \times 10^{19}$ cm$^{-3}$ of indigo [8]. The data have been obtained by calorimetry; the fitting procedure with eqs. 54 and 55 yields $a \simeq 1.8$ for the whole temperature range explored.

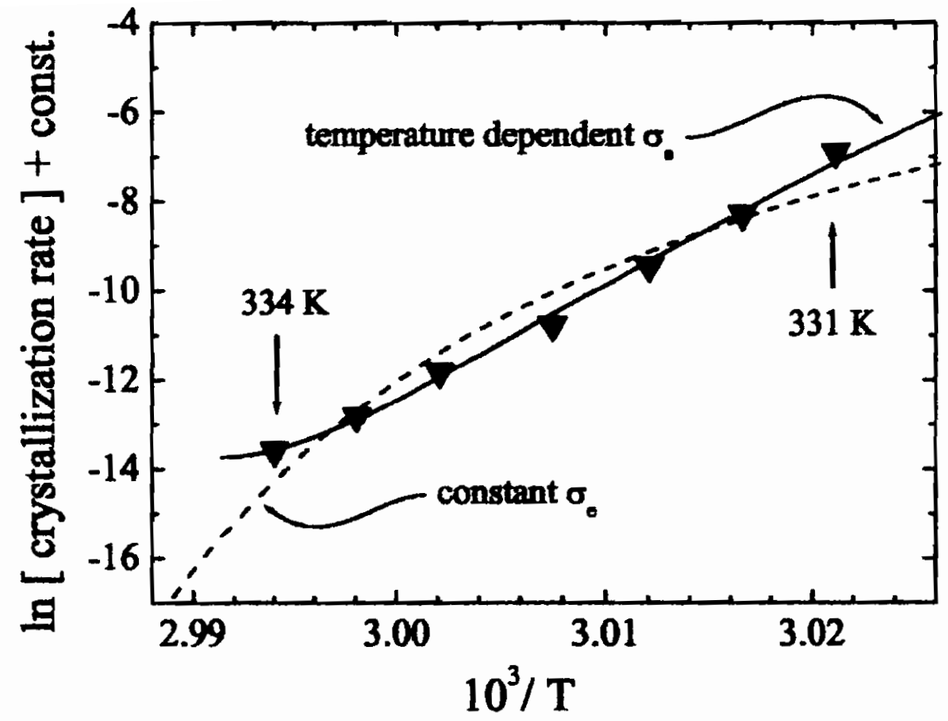



Figure 10: Logarithm of the crystallization rate (approximately $\ln v^2$) as a function of $T^{-1}$ in nucleated PEO; the symbols are the experimental data as obtained from a calorimetric analysis [7, 8], the dashed line best fits the data with a constant $\sigma_e$ while the solid line is the best fitting obtained by using eq. 55.

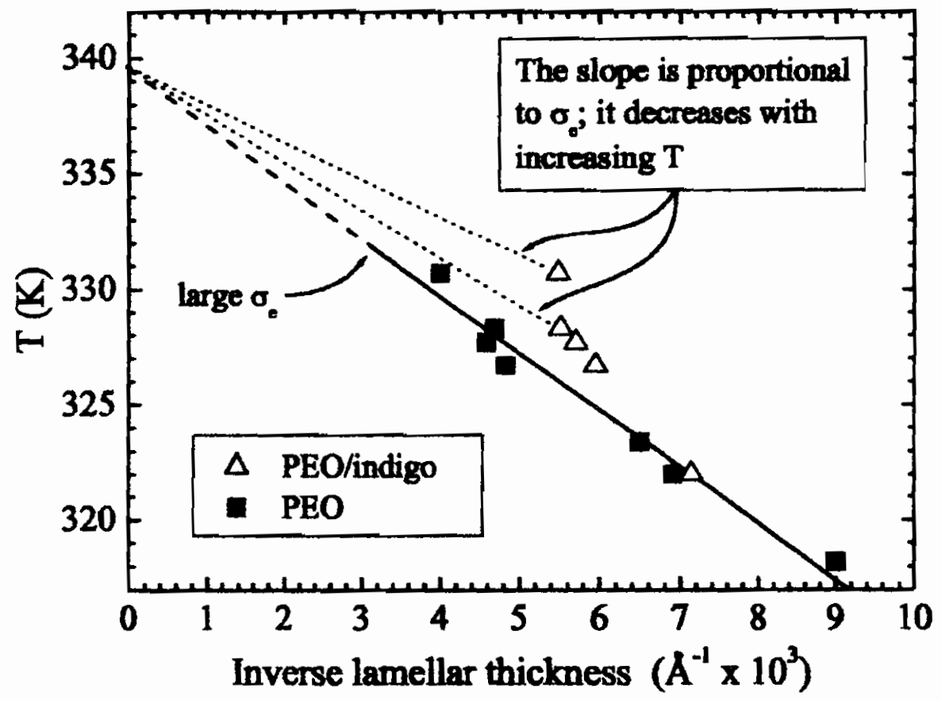



Figure 11: Gibbs-Thomson plot for the same system of Fig. 10; the data are worked out fron SAXS analysis.

Figure 11 is a Gibbs-Thomson plot where the data relative to PEO samples *with* and *without* indigo are compared [9]. The nucleant concentration is the same as that of the calorimetric measurements. The value of $a$ for the four highest temperature data ranges in the interval $1.6 < a < 2$.

These data are in fairly good agreement, and demonstrate the role of the basal surface tension in determining both growth kinetics and lamellar morphology during lateral growth. We limit ourselves to point out this consistency, although its physical basis has been described in more detail by means of a lattice model of the basal interfacial region [9, 10]. This model was able to explain the basic mechanisms responsible of the decrease in $\sigma_e$, and to predict the value of $a$ at least for PEO and polyethylene.

## 2.5 Confinement-induced vitrification in PET

### 2.5.1 introduction

An important point which has been left aside in the Adam-Gibbs theory is about the characteristic minimal dimension $n^*$ of a CRR around $T_g$. In the context of that theory, its determination relies upon assuming how large the minimal configurational entropy $s_c^*$ of a CRR should be (see eq. 43). Although reasonable, the choice $s_c^* = k_B \ln 2$ is arbitrary and not without problems because it leads to underestimates of $n^*$. Besides, also the fact that the chemical potential barrier $\delta\mu$ can be assumed to be a constant, was not given a thorough justification in the original paper. The experimental determination of the CRR dimensions is thus of prominent importance in the framework of the glass transition theories, though not at all easy. A possibility is represented by studying the glass transition process in confining geometries. Following the general idea drawn by Adam and Gibbs, a progressive decrease of the temperature would cause the extension of a CRR to increase until it becomes so close to the confining volume that structural arrest occurs. This would give a measure of the CRR volume at $T_g$. The method, however, usually has several shortcomings, not last the unpredictable

effects induced by the confining system on the sample through the contact surface.

Some progress, however, can be made in the case of PET (and possibly with other polymers as well) where the vitrification of the interlamellar amorphous component caused by a progressive decrease of the confining volume is observable. This is in fact accomplished by driving crystallization in particular conditions.

### 2.5.2 Cold drawn PET and the crystal thickening mode

The crystallization mode we have considered in the previous subsection refers to the situation where the transition proceeds by nucleation from a metastable undercooled melt. Phase separation, however, takes place by spinodal decomposition if the melt is "quenched" to a state where it is unstable (by quench we mean either a true, fast and deep temperature drop from the condition of an equilibrium melt, or an alternative procedure by which the final state is unstable anyway).

Metastability is characterized by positive values of the second derivative of the free energy $G$ of a system with respect to the order parameter $\phi$.

In the opposite case, the system is unstable with respect to the development of periodic order parameter patterns, and a non-activated (spontaneous) transition takes place. This mechanism of phase separation is better explained at the end of this section.

Crystallization of PET driven at very deep undercoolings shows characteristic features suggesting that a spinodal decomposition process underlies the overall transition to the crystalline state. The clearest evidence of this fact is that in these conditions, a periodic density modulation is observed (by SAXS) *before* any WAXS pattern is detected [11, 12].

From the practical point of view, PET is first quenched to a glassy state (i.e. below 70 °C) from the equilibrium melt. From this state of isotropic glass, PET is then re-heated to the annealing temperature $T_a$ where crystallization is planned to be driven.

The data relating to un-oriented PET crystallized at $T_a = 80$ and 120 °C, reported in the Gibbs-Thomson plot of Fig. 12 (refs. [11, 13]), are close to the line intercepting the ordinates at $T_m^\infty$. Crystal growth by thickening is active in the $T_a = 80$ °C crystallizing sample, as shown explicitly in the figure. However, the entanglements, or some other basal interfacial mechanism limiting the configurational rearrangement of the chains, eventually hinder further thickening, so that the Gibbs-Thomson behavior is followed at the end.

On the other hand, a very interesting situation is encountered when crystallization is driven in PET samples where the chains were previously oriented along a given direction. Such a non-isotropic chain configuration is not a natural condition in general, but it can be easily obtained by mechanically drawing the isotropic glass (cold-drawing).

In practice these samples are obtained from PET amorphous sheets about 0.5 mm thick; after drawing, the final sample length along the drawing direction increases naturally to about four times the initial one. After chain orientation the material is still in a glassy state.

As the temperature is raised above the glass transition temperature (which now sets in at about 60 °C, as shown in the thermogram of Fig. 3), the chain configurational degrees of freedom de-freeze in a highly non-equilibrium condition, and crystallization suddenly takes place by spinodal decomposition. Most important is the fact that the *crystal thickening* growth mode becomes now active. This is explicitly reported in the figure for the $T_a = 80$ °C un-oriented sample; for oriented PET, however, this is in fact strongly suggested by the vitrification process of the interlamellar regions as explained in the following. Figure 12 shows the Gibbs-Thomson plot for cold drawn and unoriented PET samples annealed at different temperatures $T_a$; the inset shows a schematic of the lamellar orientation with respect to the average chain direction. We note the following:

- The Gibbs-Thomson relation is fulfilled by *cold drawn* samples annealed at $T_a \geq 160$ °C; the $l_0^{-1} \to 0$ ordinate intercept

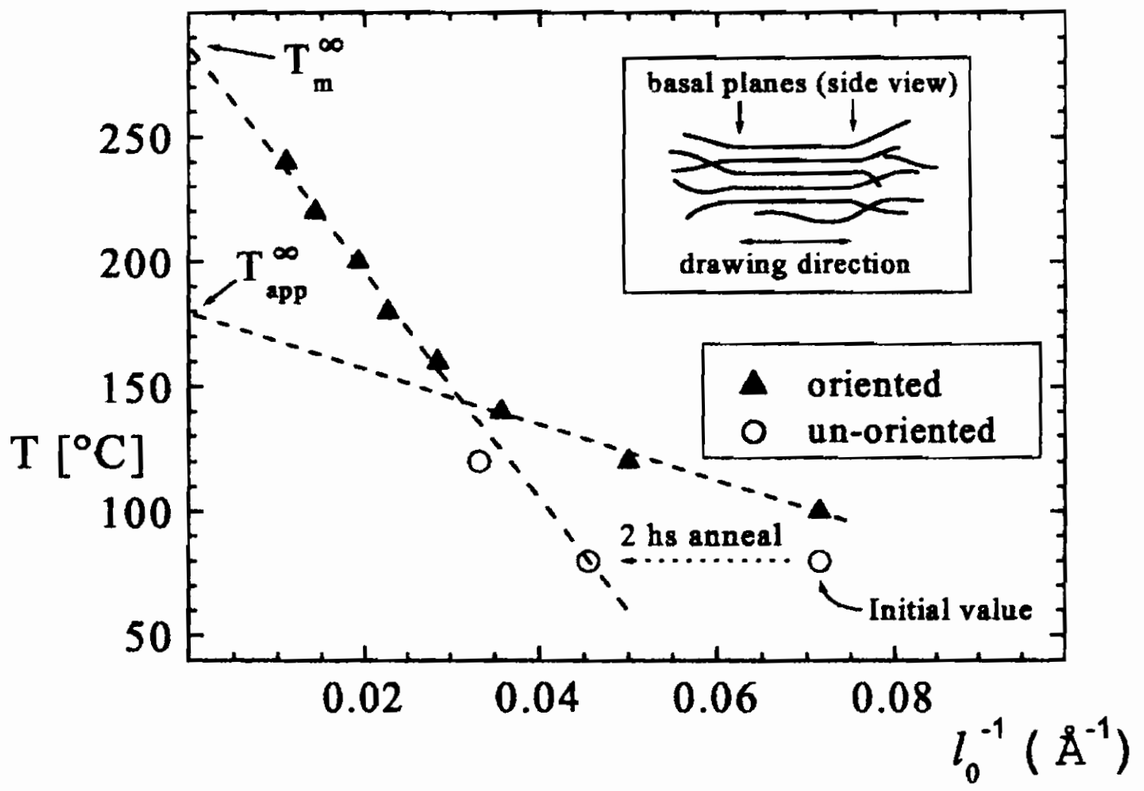



Figure 12: Gibbs-Thomson plot for cold-drawn (triangles) and un-oriented (open circles) PET samples. $T_m^\infty$ is the equilibrium melting point; $T_{app}^\infty$ is an *apparent* equilibrium melting point.

of the linear best fit to these points indicates an equilibrium melting at $T_m^\infty \simeq 285$ °C, which is in agreement with the value of PET known from the literature.

- The data of *un-oriented* samples annealed at temperatures below $T_a = 140$ °C (see refs. [11, 13] are very close to the linear best fit to the high-$T_a$ cold-drawn PET data.

- Below $T_a \approx 150$ °C the lamellar thickness of the cold-drawn samples is systematically smaller than expected from the extrapolation to low $T_a$'s of the above best fit behavior. These data, on the other hand, approximately fit a straight line intercepting the ordinate at an apparent equilibrium melting point $T_{app}^\infty \simeq 180$ °C.

From these observations we may already put forward some important questions:

1. Why do the un-oriented and (high-$T_a$) cold-drawn PET data

both fit *the same* Gibbs-Thomson behavior and extrapolate to the correct $T_m^\infty$? Would it mean that the entanglement effects are insensitive to chain orientation, or that some dominant mechanism other than entanglements contributes to hinder lamellar thickening?

2. Why lamellae formed in low-$T_a$ annealed oriented PET cannot thicken to values which the un-oriented samples are able to reach?

These questions cannot be given a conclusive answer at present, and are still an open problem in this area of research. Some light however can be shed upon by DMTA.

Figure 13 shows the relevant mechanical response of cold-drawn PET samples previously annealed at different temperatures. Instead of varying the frequency, the temperature is increased continuously from room temperature up to well above the annealing temperature. Consider the $T_a = 100$ °C annealed sample first. This sample is known by SAXS and WAXS to be partially crystalline, with a lamellar stack periodicity of about 11 nm and a lamellar thickness of about 1.4 nm [14].

Upon progressive heating the glass transition region is approached, and the process which sets in first is the de-freezing of the configurational degrees of freedom (at about 90 °C). The kinetic character of this relaxational transition is pointed out by the frequency dependence of the corresponding loss peak onset. Then, at about $T_r \approx 120$ °C the slope of the storage modulus undergoes a sudden change, which reveals that the progressive increase of the chain mobility is contrasted by an increase in the rigidity of the material as a whole. The latter is a consequence of a re-crystallization process: once the temperature has raised enough above the annealing temperature, the crystallites become unstable and the polymer naturally tends to re-crystallize into thicker lamellae. It is worth noticing, however, that the storage modulus still depends significantly on the frequency even after the thickening process set in, which means that some glassy domain is still present. Note that

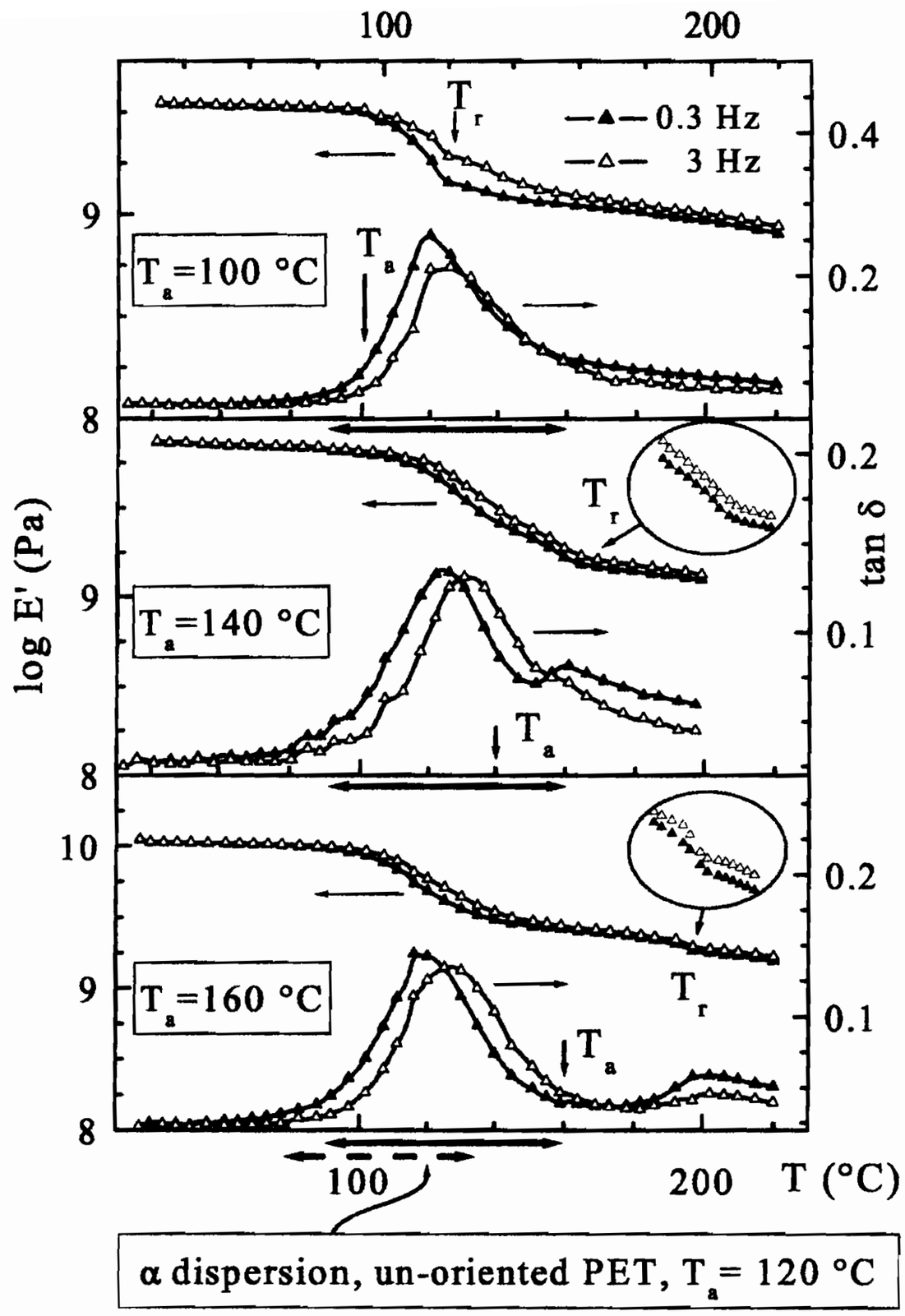



Figure 13: DMTA scans at a heating rate of 2 °C/min on cold-drawn PET annealed at different temperatures $T_a$. Note that the $\alpha$ dispersion region (highlighted by a double arrow just below the abscissas) is the same for all of them, ranging between $\sim$ 90 °C and $\sim$ 160 °C. The annealing temperatures are indicated by an arrow so to point out their location with respect to the $\alpha$ dispersion. When heating the samples, a pre-melting/recrystallization process is observable at $T_r \approx T_a + 20$ °C, independent of $T_a$, and thus, of whether the interlamellar regions are glassy or not. For a comparison, the $\alpha$ dispersion region of an un-oriented PET sample annealed at $T_a = 120$ °C is also reported; note that in this case $T_a$ falls outside the $\alpha$ dispersion (dashed double arrow).

the original annealing temperature falls well within the $\alpha$ dispersion region (i.e. the main loss peak).

All this means that after annealing, before rising the temperature for the DMTA scan, the sample was initially characterized by the presence of glassy and crystalline domains, arranged in regular stacks (otherwise they wouldn't have been observable by SAXS). However, regular lamellar stacks cannot form within a glassy material, thus the crystalline domains must have formed before the amorphous viscoelastic regions had turned to a glass. Thus, as the annealing process starts, spinodal decomposition sets in and very thin lamellae form initially, following the periodicity imposed by the unstable order parameter modes (which define the stacks' long period). Afterwards the crystalline domains tend to thicken, but in doing this the interlamellar regions become ever more constrained, so that eventually structural arrest intervenes. The final result is that the crystalline domains stop thickening due to the vitrification of the interlamellar regions. *The same holds for all samples whose annealing temperature falls within the $\alpha$ dispersion region.* (Note however that the re-crystallization process is *always* observed at about $T_r \approx T_a + 20$ °C, for any value of the annealing temperature, i.e. independent of the fact that the interlamellar region is a glass or not (see the DMTA results and the DSC scans of Fig. 14 as well).

We are now in a very advantageous condition with regards to the study of the glass transition process in confined geometries. Indeed there are three points which are worth considering:

- The vitrifying and the confining materials are of the same nature.
- The transition to a glass likely takes place gradually; in this case a sort of a crystal/glass coexistence condition sets in, so the morphology of the crystal domains would possibly carry indirect information about the embedded glass.
- The difference $T_r - T_a$ is a constant, independent of the fact

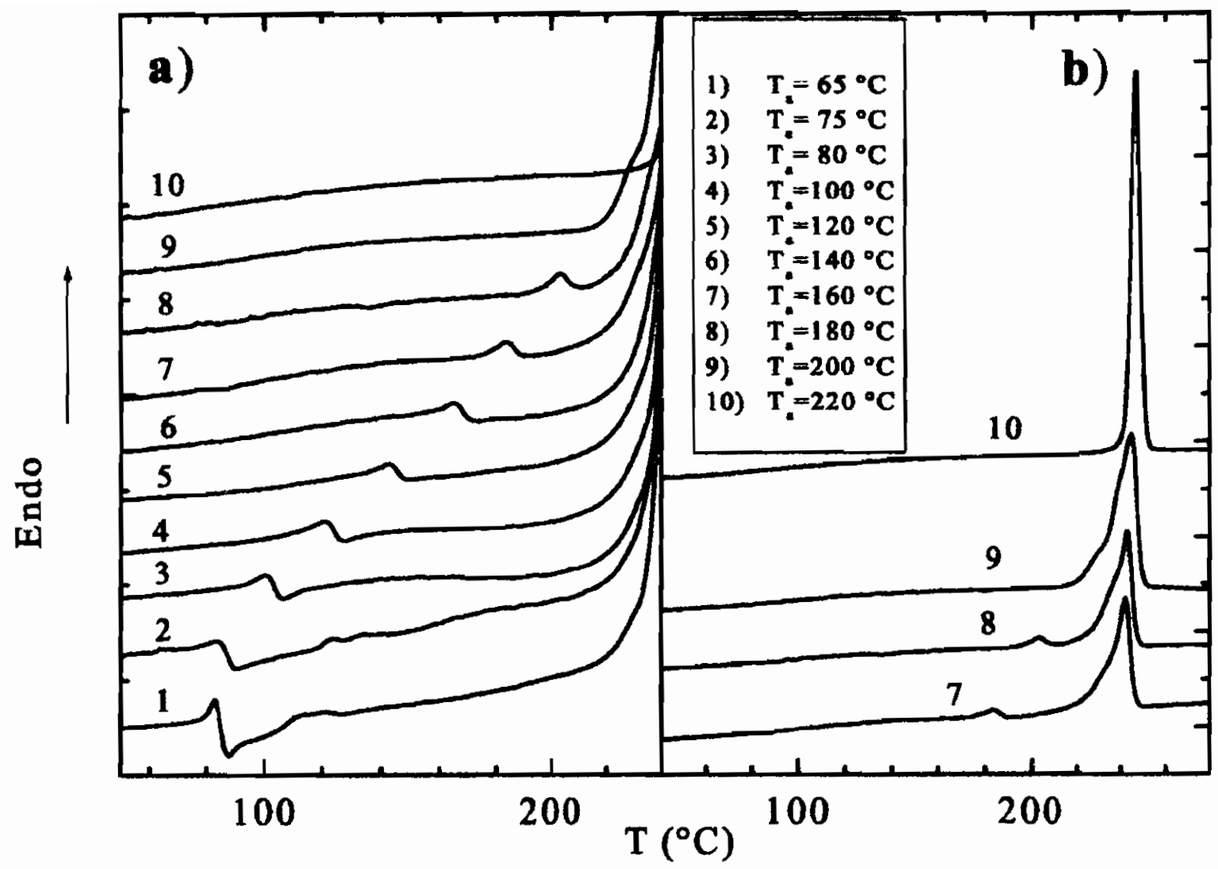



Figure 14: DSC scans at a heating rate of 40 °C/min on cold-drawn PET samples annealed at different temperatures. Note that all samples show an endothermal peak slightly above the corresponding annealing temperature. This endothermal process corresponds to an increase of the configurational motion of the chains as a precursor to further crystallization. For the low-$T_a$ annealed samples, the crystallization process is pointed out rather clearly by an exothermal component just following the maximum of the precursor process. For $T_a \geq 160\,°C$ the crystallization process is not visible anymore, but it actually takes place as can be seen by DMTA (see the inset of the lowest panel in Fig. 13).

that the interlamellar regions are either in a glassy state or not.

Now, two more questions arise:

1. Is it possible to formulate a balance condition for the low-$T_a$ annealed oriented PET?

2. Why do the oriented amorphous domains vitrify, whereas it is not so for the un-oriented domains?

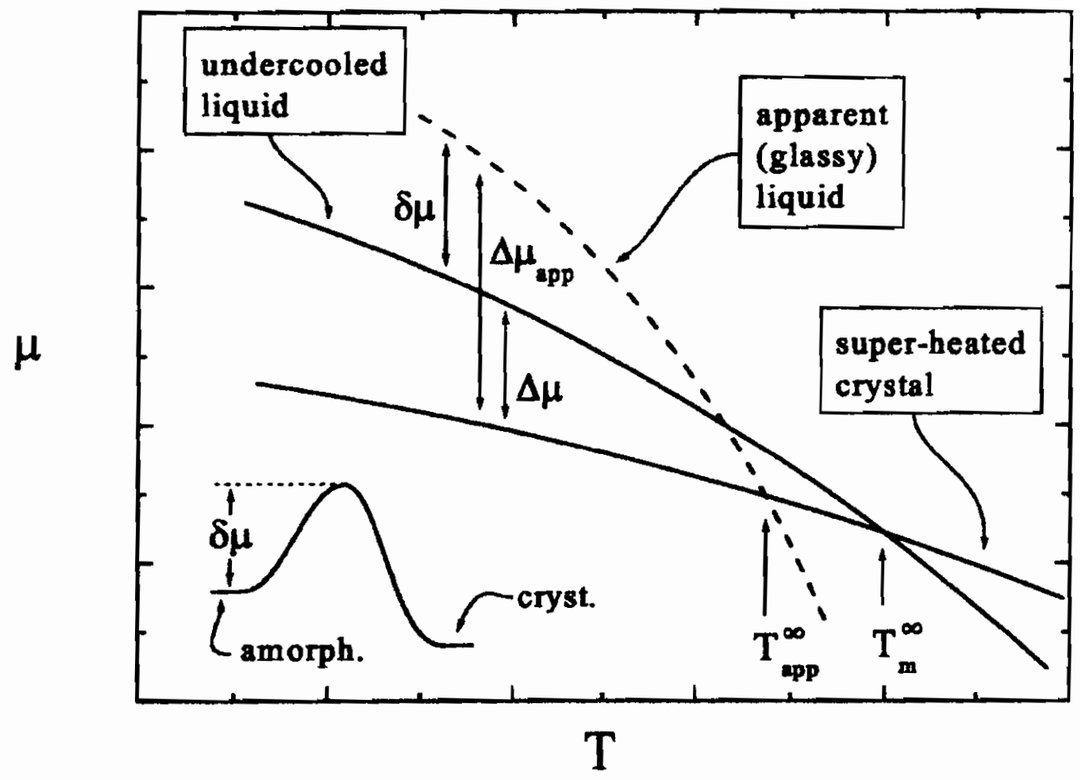



Figure 15: Liquid and solid branches of the chemical potential $\mu$ as functions of the temperature (solid lines). The dashed line represents an hypothetical "liquid" branch which would mimic the effect of the glass.

About question 1), one may in fact relate the linear intercept to $T^{\infty}_{app}$ of the Gibbs-Thomson plot, to the temperature dependence of a hypothetical chemical potential associated to the glass (as in Fig 15). The gradual arrest of the thickening process would be viewed as a consequence of an increase of the *apparent* chemical potential of the amorphous phase (see eq. 51).

At first sight this may appear as a hasty speculation, but after thinking a little about, this point of view seems to fit some of the general ideas about the glass transition we saw above. Indeed, in introducing the fundamentals about the trend of thermodynamic processes in general, we saw that the establishing of an internal constraint has the effect of rising some thermodynamic potential. This has been already formalized by Adam and Gibbs who with the aim of describing de-frozen units in the amorphous domain, simply dropped from the partition function all the Boltzmann factors associated to frozen-in configurational states (i.e. those states from

which a CRR cannot undergo any configurational transition). As a consequence, the need to overcome a chemical potential barrier $\delta\mu$ for structural rearrangement to occur, results as an effect of a reduced number of "active" states. The representation suggested above shows similar features: structural arrest is equivalent to the appearance of internal constraints and the difference $\Delta\mu_{app} - \Delta\mu$ plays the role of the barrier $\delta\mu$ introduced by Adam and Gibbs (indeed, the chain segments at the basal boundary which would be able to "enter" the crystalline phase, should already be around the top of the barrier).

In order to give further support to this hypothesis, we are now left with the problem of determining the partition function of the amorphous phase, and see how this idea fits with the concepts of statistical mechanics. In this respect there is one more point which is worth considering, namely, question 2 above. Here again the experiment comes for a help.

Consider the PET samples annealed at $T_a = 120\,°\mathrm{C}$; the unoriented sample follows the Gibbs-Thomson relation, while the oriented one does not. On the other hand, SAXS analysis reveals that the thickness of the interlamellar regions are practically the same in both the cases, i.e. $\sim 5.5$ nm. Now, during the lamellar thickening process, two opposing basal planes embedding a given amorphous region cease to move one towards the other when the interlamellar defects "pushed away" from one side impinge onto the other. In other words, when some correlation length starts being comparable with the interlamellar thickness.

Since defect-free regions are more dense than defect-rich ones, we should check whether, and to what extent, the density-density correlation length is affected by chain orientation. Moreover, if a difference is found (as it will be), then it is important to write down the partition function as an expansion in normal modes of the density fluctuations, because this will allow to address the connection between chemical potential and fluctuating modes more directly.

To this aim we shall resort to the gaussian chain model of a polymeric chain, because it is able to catch a number of character-

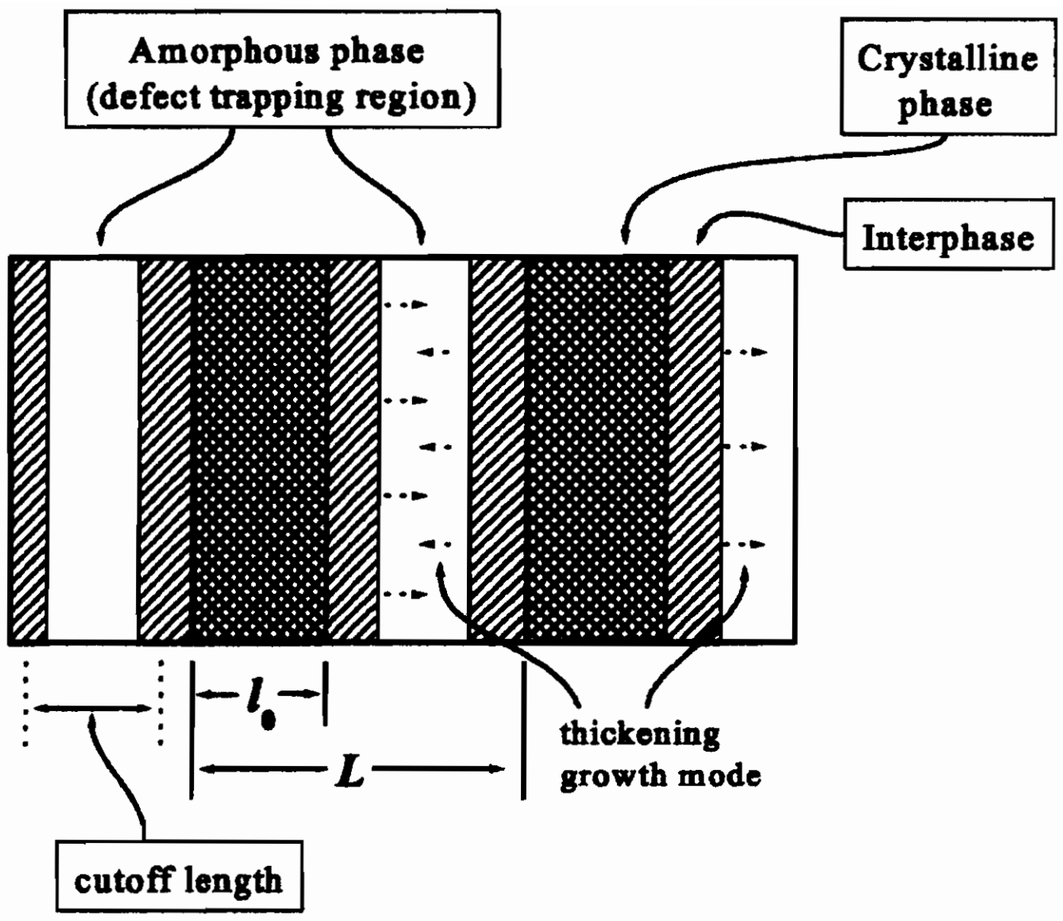



Figure 16: Crystalline/Interfacial/Amorphous layers in a stack. $L$ is the long period and $l_0$ is the lamellar thickness; the cutoff length is roughly $\sim (L - l_0)$. Vitrification should occur when the "cutoff length" becomes comparable with the density-density correlation length.

istic features of polymer statistics, still remaining simple enough to allow for some analytical calculation. This model will be introduced in next subsection; some similarity with the formalism of quantum mechanics will be also pointed out.

### 2.5.3 The gaussian chain model

The simplest way to describe the conformation of polymers is to discretise the space into a lattice. We shall consider the chains to be made of units (usually called *segments*) which are thought to be accommodated at the sites of a cubic lattice; these segments are then connected by *bonds* to form chains (see Fig. 17). A segment does not in general coincide with just one monomeric unit, but may often be larger. The main features of the chain conformation statistics can be catched by assuming that the directions of subsequent

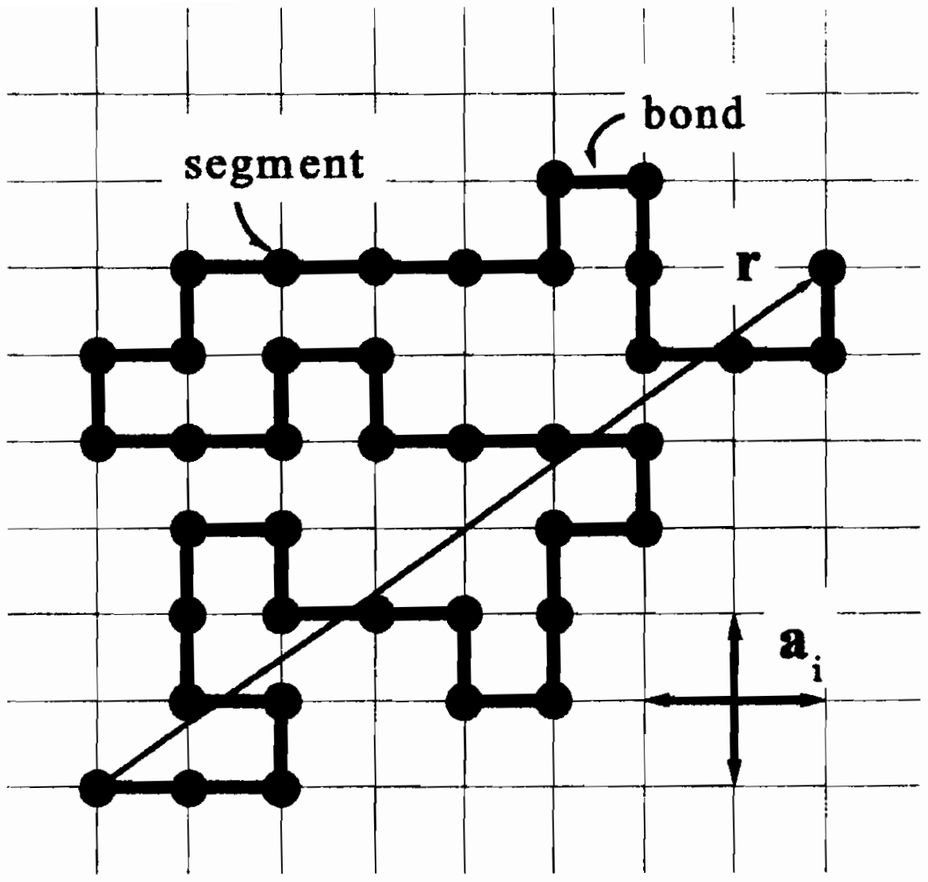



Figure 17: A chain made of segments into a cubic lattice.

bonds are not mutually correlated (freely jointed chain), and that multiple occupation of any lattice site is allowed (ideal or phantom chain).

Having fixed one of the two chain ends to a given site (the origin), we seek for the probability $P(\mathbf{r};n)$ that the other end of a chain made of $n$ segments is at point $\mathbf{r}$. Due to connectivity, the $(n-1)$th segment can only be in one of the nearest neighboring sites about $\mathbf{r}$ with equal probability $z^{-1}$ (freely jointed approximation), $z=6$ being the coordination number in a cubic lattice:

$$P(\mathbf{r},n)=\frac{1}{z}\sum_{i=1}^{z}P(\mathbf{r}-\mathbf{a}_i,n-1)\,, \tag{57}$$

where the $\{\mathbf{a}_i\}$ are the principal lattice vectors. For very long chains, i.e. $|\mathbf{r}|\gg|\mathbf{a}_i|$ and $n\gg 1$, the r.h.s. of eq. 57 can be expanded to yield:

$$P(\mathbf{r}-\mathbf{a}_i,n-1)\simeq P(\mathbf{r};n)-\frac{\partial P}{\partial n}-\mathbf{a}_i\cdot\nabla P+\frac{1}{2}\mathbf{a}_i\,\mathbf{a}_i:\nabla\nabla P\,. \tag{58}$$

Since $\sum_i \mathbf{a}_i = 0$ and $z^{-1}\sum_i a_{i\mu}a_{i\nu} = \delta_{\mu\nu}a^2/6$, with $a \equiv |\mathbf{a}_i|$, we find immediately

$$\frac{\partial P}{\partial n} = \frac{a^2}{6}\nabla^2 P\,. \tag{59}$$

Last one is a very well known equation in the theory of the random walk, and is also of the same form an *imaginary time* Schrödinger equation for a free particle. The solution is the gaussian distribution:

$$P(\mathbf{r}, n; \mathbf{r}_0, 0) = \left(\frac{3}{2\pi a^2 n}\right)^{3/2} \exp\left\{-\frac{3|\mathbf{r}-\mathbf{r}_0|^2}{2a^2 n}\right\}, \tag{60}$$

which describes an ideal polymer chain in isotropic conditions (the initial coordinates of the chain have been explicitly indicated). A fundamental property of the solution expresses the composition of probabilities:

$$P(\mathbf{r}, n; \mathbf{r}_0, 0) = \int d^3\mathbf{w}\, P(\mathbf{r}, n; \mathbf{w}, \tau)\, P(\mathbf{w}, \tau; \mathbf{r}_0, 0)\,; \quad \forall\, \tau < n\,. \tag{61}$$

We will generalize this result to the continuum, so we allow for this fractionation to be repeated at will; dividing the contour length in intervals of equal length $\delta\tau$, we find

$$P(\mathbf{r}, n; \mathbf{r}_0, 0) = \frac{1}{[2\pi a^2 \delta\tau/3]^{3n/2}} \int \prod_{i=1}^{n-1} d^3\mathbf{r}_i \exp\left\{-\frac{3}{2\,a^2}\sum_{i=0}^{n-1}\frac{(\mathbf{r}_{i+1}-\mathbf{r}_i)^2}{\delta\tau}\right\}, \tag{62}$$

which in the limit of an infinite partition is in fact the functional integral with respect to all possible paths $\mathbf{r}(\tau)$:

$$P(\mathbf{r}, n; \mathbf{r}_0, 0) = \int_{\mathbf{r}_0}^{\mathbf{r}} \mathcal{D}\mathbf{r}(\tau) \exp\left\{-\frac{3}{2\,a^2}\int_0^n d\tau\, \dot{\mathbf{r}}^2\right\}, \tag{63}$$

the dot standing for the derivative with respect to $\tau$: $\dot{\mathbf{r}} \equiv \delta\mathbf{r}/\delta\tau$.

Equation 61 says that the probability that a chain starts at $\mathbf{r}_0$ and ends at $\mathbf{r}$ after $n$ steps, can be obtained from the combined probability that the chain starts at $\mathbf{r}_0$, passes through $\mathbf{w}$ at step $\tau$,

and ends at $\mathbf{r}$; for the overall probability to be obtained, however, we have to sum over all possible intermediate points $\mathbf{w}$. If a field is also present, so that an energy $U(\mathbf{w})$ is associated to a chain segment placed at $\mathbf{w}$, then the possible contributions to the integral must be weighted by a Boltzmann factor $\exp\{-U(\mathbf{w})/k_BT\}$. A straightforward generalization for an infinite partitioning of the path thus leads to

$$P(\mathbf{r}, n; \mathbf{r_0}, 0) \sim \int_{\mathbf{r_0}}^{\mathbf{r}} \mathcal{D}\mathbf{r}(\tau) \exp\left\{-\int_0^n d\tau \left[\frac{3}{2\,a^2}\dot{\mathbf{r}}^2 + \frac{U(\mathbf{r})}{k_BT}\right]\right\} . \quad (64)$$

This form of the probability suggests that in fact it may be interpreted as the *partition function of a single chain* with the end points fixed at $\mathbf{r}$ and $\mathbf{r_0}$, the functional integral accounting for all conformational fluctuations. A well defined chain configuration $\{\mathbf{r}(\tau)\}$ is thus assigned a Boltzmann factor, customarily called the *segment distribution*,

$$\Psi\{\mathbf{r}(\tau)\} = A \exp\left\{-\int_0^n d\tau \left[\frac{3}{2\,a^2}\dot{\mathbf{r}}^2 + \frac{U(\mathbf{r})}{k_BT}\right]\right\} , \quad (65)$$

with $A$ a suitable normalization factor (do not confuse this $\Psi$ with the general thermodynamic potential in eq. 22).

### 2.5.4 Density fluctuation mode representation of an oriented melt

In order to describe a polymer melt, we introduce an excluded volume mean field interaction with an energy term

$$U_e = \frac{v}{2}\sum_{i,j}\int_0^n d\tau \int_0^n d\tau' \delta[\mathbf{r}_i(\tau) - \mathbf{r}_j(\tau')] , \quad (66)$$

with $i$ and $j$ labelling the different chains; $v$ is the excluded volume interaction parameter [15]. Concerning the "elastic" (entropic) component of the energy, i.e. the term involving $\dot{\mathbf{r}}$ in $\Psi$, the chains are treated as ideal. Since we also want to generalize the theory to

the case where the chains have a mean orientation along a preferred direction, the overall elastic energy will be taken in the form

$$U_0 = k_B T \frac{3}{2\,a^2} \sum_j \int_0^n d\tau \left[ \dot{\mathbf{r}}_{\perp j}^2 + \kappa^{-1} \dot{\mathbf{r}}_{\parallel j}^2 \right] , \tag{67}$$

where $\kappa = 1$ for an isotropic melt, and $\kappa > 1$ for chains oriented (stretched) along the "parallel" ($\parallel$) direction.

From the overall chain distribution

$$\Psi_t\{\mathbf{r}_i(\tau)\} \propto \exp\{-(U_0 + U_e)/k_B T\} \tag{68}$$

the partition function might be obtained by a functional integration over all possible conformations $\mathbf{r}_i(\tau)$, thus allowing for all configurational fluctuations to be accounted for. However, we are interested to a *density mode expansion* of the partition function, because we saw above that the density-density fluctuations should be relevant in characterizing the behavior of the system when the structural arrest is approached through confinement. We then first project $\Psi_t\{\mathbf{r}_i(\tau)\}$ onto the space of the density mode amplitudes.

To this aim, consider the overall density $c(\mathbf{r})$ together with the associated Fourier transform $c_{\mathbf{k}}$:

$$c(\mathbf{r}) = \sum_j \int_0^n d\tau\, \delta[\,\mathbf{r} - \mathbf{r}_j(\tau)] \tag{69}$$

$$c_{\mathbf{k}} = \frac{1}{V} \sum_j \int_0^n d\tau\, e^{i\mathbf{k}\cdot\,\mathbf{r}_j(\tau)} \tag{70}$$

[note that $c_{-\mathbf{k}} = c_{\mathbf{k}}^*$ for $c(\mathbf{r})$ to be real]. Then the segment distribution with respect to the possible sets $\{c_{\mathbf{k}}\}$ can be obtained as

$$\Psi_t\{c_{\mathbf{k}}\} \propto \prod_{\mathbf{k}>0} \int \prod_i \delta\mathbf{r}_i \Psi_t\{\mathbf{r}_i(\tau)\}\, \delta \left[ c_{\mathbf{k}} - \frac{1}{V} \sum_i \int_0^n d\tau\, e^{i\mathbf{k}\cdot\,\mathbf{r}_i(\tau)} \right] , \tag{71}$$

where $\mathbf{k} > 0$ means that each mode must be accounted for only once (remind that $c_{\mathbf{k}} = c^*_{-\mathbf{k}}$). From the expansion of the delta function, $\delta(\mathbf{r}) = V^{-1} \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}}$, we immediately see that the excluded volume interaction can be expressed as a quadratic form in the mode amplitude:

$$U_e\{c_{\mathbf{k}}\} = k_B T \frac{vV}{2} \sum_{\mathbf{k}} c_{\mathbf{k}} c_{-\mathbf{k}} \,. \tag{72}$$

But with the entropic contribution $U_0$ things are not as simple; by the way, we seek for a quadratic approximation as well:

$$U_0\{c_{\mathbf{k}}\} \simeq k_B T \sum_{\mathbf{k}} u_{\mathbf{k}}\, c_{\mathbf{k}} c_{-\mathbf{k}} \,, \tag{73}$$

which is expected to be reliable in the case of small density fluctuations.

The coefficients $u_{\mathbf{k}}$ are directly related to the structure factor of the ideal chain and are easily calculated below. Due to the independence of the modes with different $\mathbf{k}$, the chain distribution $\Psi_t\{c_{\mathbf{k}}\}$ reduces to the product of single mode *gaussian* distributions. In the case that $U_e \equiv 0$, the coefficients $u_{\mathbf{k}}$ are thus related to the average amplitude of the mode by

$$u_{\mathbf{k}} = \frac{1}{2\langle c_{\mathbf{k}} c_{-\mathbf{k}} \rangle_0} \,, \tag{74}$$

where the subscript "0" indicates that the average is calculated for *ideal* chains. Because of the statistical independence of different chains in the absence of excluded volume interactions,

$$\begin{aligned} \langle c_{\mathbf{k}} c_{-\mathbf{k}} \rangle_0 &= \frac{1}{V^2} \sum_{i,j} \int_0^n d\tau \int_0^n d\tau' \, \langle e^{i\,\mathbf{k}\cdot(\,\mathbf{r}_i - \mathbf{w}_j)} \rangle_0 \\ &= \frac{M}{V^2} \int_0^n d\tau \int_0^n d\tau' \int d^3\mathbf{r}\, d^3\mathbf{w}\, P(\mathbf{r},\tau;\mathbf{w},\tau')\, e^{i\,\mathbf{k}\cdot(\,\mathbf{r}-\mathbf{w})} \,, \end{aligned} \tag{75}$$

where $M$ is the number of chains in the volume $V$. This gaussian integral can be readily calculated and we finally get

$$\langle c_{\mathbf{k}} c_{-\mathbf{k}} \rangle_0 = \frac{c\, S_{\mathbf{k}}}{V} \,, \tag{76}$$

with $c \equiv Mn/V$ the average segment (number) concentration, and

$$S_{\mathbf{k}} \equiv \frac{2n}{x^2}\left(e^{-x} - 1 + x\right), \quad x \equiv \frac{n\,a^2}{6}(\,\mathbf{k}_\perp^2 + \kappa\, k_\parallel^2) \tag{77}$$

the ideal chain structure factor.

For our scopes, the large $x$ limit of the structure factor is relevant, so we approximate $S_{\mathbf{k}} \approx 12\,[\,a^2(\,\mathbf{k}_\perp^2 + \kappa\, k_\parallel^2)]^{-1}$ and the final expression of the segment distribution takes the form

$$\Psi_t\{c_{\mathbf{k}}\} \propto \exp\left\{-\frac{V\,a^2}{24c}\sum_{\mathbf{k}}(\,\mathbf{k}_\perp^2 + \kappa\, k_\parallel^2 + \xi^{-2})\,c_{\mathbf{k}}c_{-\mathbf{k}}\right\}, \tag{78}$$

where the parameter $\xi \equiv (8acv)^{-1/2}$ is called the *correlation length.* So, the excluded volume interaction tends to depress the average density fluctuation mode amplitude, since

$$\langle c_{\mathbf{k}}c_{-\mathbf{k}}\rangle = 12\,c/[\,V\,a^2(\,\mathbf{k}_\perp^2 + \kappa\, k_\parallel^2 + \xi^{-2})] < \langle c_{\mathbf{k}}c_{-\mathbf{k}}\rangle_0\,. \tag{79}$$

The knowledge of $\Psi_t\{c_{\mathbf{k}}\}$ in this form is the basis for the forthcoming discussion.

Before concluding this subsection, it may be instructive to show why $\xi$ has been called "correlation length". Suppose that the density at some point (the origin) deviates with respect to the mean value $c$, and ask how far from the origin this perturbation decays completely. To this aim one is led to calculate the quantity

$$\langle \delta c^2\rangle \equiv \langle c(\mathbf{r})\,c(\mathbf{0})\rangle - c^2 = \frac{V}{8\,\pi^3}\int d^3\mathbf{k}\,\langle c_{\mathbf{k}}c_{-\mathbf{k}}\rangle\,e^{-i\,\mathbf{k}\cdot\,\mathbf{r}}\,. \tag{80}$$

The calculation of this integral will be outlined afterwards for arbitrary $\kappa$. Here it suffices to report the result for $\kappa = 1$ (isotropic melt), which reads: $\langle \delta c^2\rangle = (3c/\pi a^2|\mathbf{r}|)\,\exp(-|\mathbf{r}|/\xi)$, and points out clearly the role of $\xi$.

### 2.5.5 Vitrification of an oriented amorphous phase

We are almost done. Once the segment distribution in the appropriate representation is available, we can find the partition function

and then see why an oriented melt should turn to a glass with better chances than the un-oriented one.

So, consider the volume $V$ of an amorphous interlamellar region, and assume it is practically a constant; the appropriate thermodynamic potential is then the Helmholtz free energy. Taking the un-oriented state as a reference, the free energy of the oriented system is

$$\Delta F \equiv F_{or} - F_u = -k_B T \ln \frac{Z_{or}}{Z_u} , \tag{81}$$

where $Z_{or}$ and $Z_u$ are the partition functions for the oriented and un-oriented amorphous systems respectively. These quantities can be expressed as integrals over the modes' amplitudes of the segment distributions: $Z \propto \int \prod_{\mathbf{k}} d\, c_{\mathbf{k}} \Psi_t\{c_{\mathbf{k}}\}$, with $\int d\, c_{\mathbf{k}} \equiv \int d\, \Re c_{\mathbf{k}} \int d\, \Im c_{\mathbf{k}}$. It is important to stress that the free energies in eq. 81 are associated to just the the configurational (density) fluctuations. No static contribution associated with the mean density (i.e. at $|\mathbf{k}| = 0$) is included.

Because of the approximation on the energy, we are led to the calculation of ordinary gaussian integrals, and the argument of the logarithm in the r.h.s. of eq. 81 finally takes the form

$$\frac{Z_{or}}{Z_u} = \prod_{\mathbf{k}>0} \frac{|\mathbf{k}|^2 + \xi^{-2}}{\mathbf{k}_\perp^2 + \kappa\, k_\parallel^2 + \xi^{-2}} . \tag{82}$$

Since $\kappa \geq 1$, the r.h.s. of last equation is an infinite product of terms all $\leq 1$, and the larger ones are roughly those for which $|\mathbf{k}|$ is small (i.e. the mode wavelength is large).

We now consider the effect of the embedding crystalline regions. The latter in fact represent a constraint for the density fluctuations of wavelength comparable or larger than the thickness of the interlamellar region. These frozen modes cannot contribute to the fluctuation free energy and must be excluded from the product (see eqs. 81 and 82).

This crude model carries with it some fundamental questions. As explained above, when the glass transition process takes place some constraints are progressively established. We can resort to

thermodynamics either before these constraints set in, or afterwards, when their characteristic relaxation times (i.e. of the constraints) become much longer than the laboratory time scale. In between, no statistical thermodynamic argument seems to be rigorously applicable in principle, and one is led to consider alternative formulations related to non-equilibrium formalisms.

Let $Z'_{or}$ and $Z'_{u}$ stand for the partition functions calculated accounting only for the un-constrained density modes; then for the new free energy difference $\Delta F'$ we find

$$\Delta F' - \Delta F = k_B T \sum_{\mathbf{k}>0}^{\mathbf{k_c}} \ln \left[ 1 + (\kappa - 1) \frac{k_{\parallel}^2}{|\mathbf{k}|^2 + \xi^{-2}} \right] , \qquad (83)$$

where the "interval" $0 < \mathbf{k} \leq \mathbf{k}_c$ is meant to include the wave vectors of all constrained (frozen) modes. The corresponding difference between the chemical potentials can be estimated from the relationship $\mu = (\partial F / \partial \mathcal{N})_{V,T}$, with $\mathcal{N} = n\,M$ the total number of segments. Note that as the interval of "excluded" modes increases, the difference between the free energies of the oriented and un-oriented regions increases too.

Now, assuming that in the case of an oriented system the chains are mainly directed perpendicular to the embedding basal planes, we make things simple without going far from the experimental evidence (see the inset of Fig. 12; for experimental data see [14]). So, the direction which we labelled with the symbol "$\parallel$" is in fact perpendicular to the basal planes, and the thickness $l_a$ of the interlamellar region establishes a cutoff for all those wave vectors for which

$$k_{\parallel} \lesssim \frac{2\,\pi}{l_a} \equiv k_c \,. \qquad (84)$$

In order to estimate the change of the chemical potential difference caused by the "quenching" of long wave "parallel" fluctuations, we turn to the continuous with the substitution $\sum_{\mathbf{k}} \rightarrow (V/8\,\pi^3) \int d^3\mathbf{k}$,

and find after straightforward calculation

$$\Delta\mu' = \Delta\mu + \frac{k_B T}{8\,\pi^2 c\,\xi^2}\left\{k_c \ln\left(\frac{\kappa\,k_c^2\xi^2+1}{k_c^2\xi^2+1}\right) + \frac{2}{\xi}\left[\frac{\arctan(k_c\xi\sqrt{\kappa})}{\sqrt{\kappa}} - \arctan(k_c\xi)\right]\right\}. \qquad (85)$$

Note that $\Delta\mu'$ is larger than $\Delta\mu$ except when $\kappa = 1$.

The point of view put forward so far, i.e. the argument based on the quenching of the modes in the partition function, is rather intuitive. The starting idea, however, involved the density-density correlation length, so it is desirable to have a glance at this quantity. As explained just at the end of previous subsection, the correlation length establishes the spatial decay rate of the density perturbations $\langle \delta c^2 \rangle$. By expressing the latter as a Fourier integral of the mode amplitudes (see eq. 80) and performing the integration over $d^2\mathbf{k}_\perp$, we are left with:

$$\langle \delta c^2 \rangle = \frac{6\,c}{\pi^2 a^2}\int_{k_c}^{+\infty} d\,k_\parallel e^{-i\,k_\parallel r_\parallel} K_0\left(|\mathbf{r}_\perp|\sqrt{\kappa k_\parallel^2 + \xi^{-2}}\right), \qquad (86)$$

where $K_0$ is a Bessel function of imaginary argument. Using the integral representation $K_\nu(ab) = 1/2(b/a)^\nu \int_0^{+\infty} dt\, t^{\nu-1} \exp\{-[a^2/t + b^2 t]/2\}$, for $k_c = 0$ we find:

$$\langle \delta c^2 \rangle = \frac{3\,c}{\pi\, a^2[\,|\mathbf{r}_\perp|^2 + r_\parallel^2/\kappa\,]^{1/2}} \exp\left\{-\frac{1}{\xi}\sqrt{|\mathbf{r}_\perp|^2 + \frac{1}{\kappa}r_\parallel^2}\right\} \qquad (87)$$

(the same result could be obtained by performing the calculation as for an isotropic system, after a suitable length rescaling along the parallel direction). The above expression says that the effective correlation length along the parallel direction is $\kappa^{1/2}\xi$, i.e. larger than in the isotropic case.

Figure 18 shows the space dependence of $\langle \delta c^2 \rangle$ for different values of $k_c$ (see eq. 86).

The quantity $\xi$ is in fact a microscopic parameter defined just below eq. 78. It is only after the integration of $\langle c_{\mathbf{k}} c_{-\mathbf{k}} \rangle$ in the $\mathbf{k}$ space

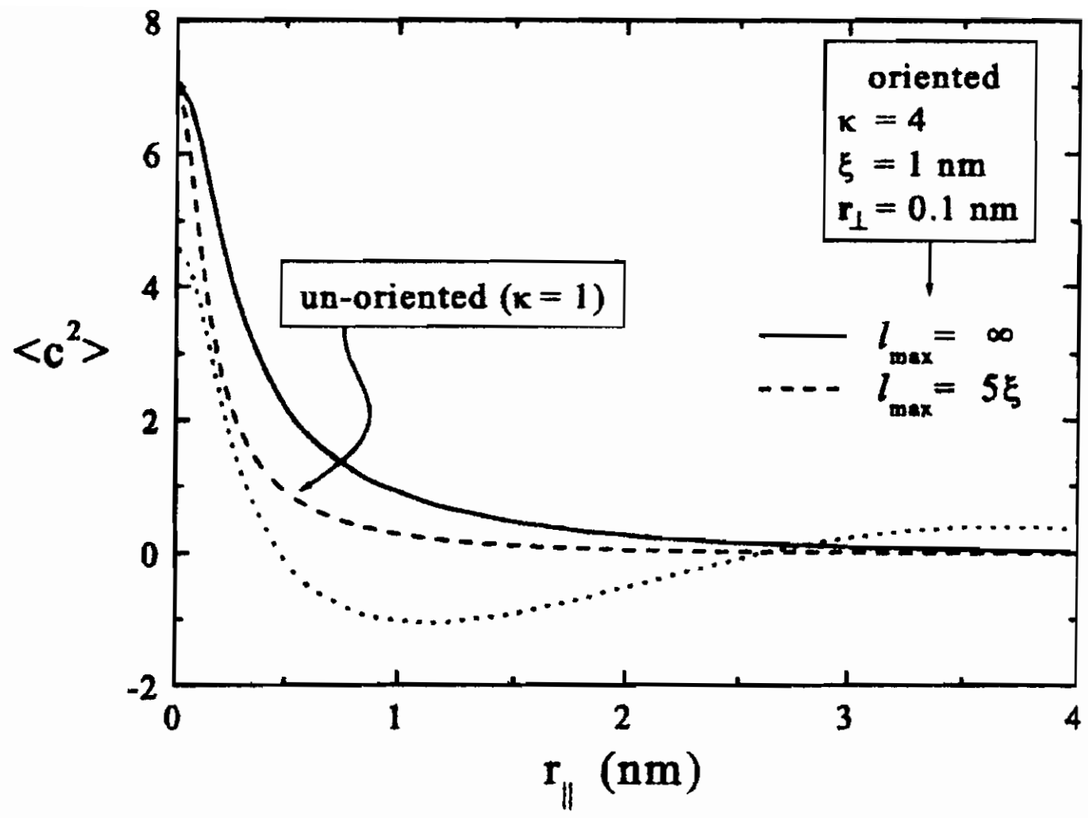



Figure 18: Density-density correlation $\langle \delta c^2 \rangle$ as a function of the distance along the drawing direction; $l_{max}$ is the cutoff length, i.e. the maximum wavelength value for the corresponding mode to be able to fluctuate.

that we see that it has the meaning of a correlation length. In other words, the correlation length naturally tends to assume the value of the microscopic parameter $\xi$ (or $\xi/\kappa^{1/2}$) *if* all modes are allowed to fluctuate. On, the other hand, when the modes with small $\mathbf{k}$ are quenched, the effective correlation length is kept smaller than the value $\xi$ (or $\kappa^{1/2}\xi$) it would assume, and we are in a condition of a constrained amorphous system, i.e. a glass.

There is a difference of course, between a glass obtained by confinement and an ordinary glass. In the latter the constraints on the long wavelength modes are established by the homogeneous phase itself, whereas in the former the constraints are a consequence of the presence of a phase heterogeneity. In this respect we also have to point out that in the scheme set forth above, all segments are treated in the same way, much as in the Adam-Gibbs theory, where the same chemical potential barrier $\delta\mu$ is assigned to each of the segments (or units) of a CRR, independent of its position with respect to the CRR boundary. Strictly speaking, we should have

carried out the calculation of the partition function for a system made of chains which are effectively anchored to the basal surfaces of the interlamellar volume. Thus the above calculation has only the value of an estimate.

But what did we really calculate? It is important at this point to make an explicit comparison with the Adam-Gibbs theory. Our main idea was to ascribe the reduced thickness of the crystalline lamellae to a change of the *apparent* bulk amorphous chemical potential as shown in Fig. 15. This way we can still describe the relationship between thermodynamic parameters and lamellar thickness, eq. 51, as the result of a *balance condition* involving the crystalline bulk, the interface and, *now*, the segments of the amorphous phase which are at the top of the chemical potential barrier $\delta\mu$ introduced by Adam and Gibbs.

It remains to discuss the dependence of $\delta\mu$ on the size of the confining region. In the Adam-Gibbs theory, the divergence of the viscosity of a glass forming liquid is ascribed to the increase of the linear dimensions [$\sim (n^*)^{1/3}$] of a CRR:

$$\eta \sim e^{n^* \delta\mu / k_B T} . \tag{88}$$

(from eqs. 42 and 44). In a sense, a CRR (i.e. $n^*$) is viewed as growing indefinitely at constant $\delta\mu$ as $T$ approaches the ideal glass transition. On the other hand, when vitrification is progressively induced by confinement, then initially $\delta\mu = 0$ and the number $n^*$ of units involved in any cooperative structural rearrangement fills a volume smaller than the confining one. As the latter decreases further, then the region occupied by the $n^*$ units starts to be constrained: the group of units involved in a structural rearrangement then turn to a CRR and $\delta\mu$ increases. So, $\delta\mu$ must depend on the linear dimensions of the confining region at the temperature where the glass transition is induced. If the condition for the liquid to become a glass is expressed through a certain threshold value $K_{Glass}$ that the exponent in eq. 88 has to reach, then we may write the proportionality

$$\delta\mu \propto \frac{1}{n^*} . \tag{89}$$

This relationship is in fact described in eq. 85 by the coefficient $k_c$ of the logarithm.

We now want to see how our calculations match the experiment, otherwise they would be useless. To this aim we consider the data of both un-oriented and cold-drawn PET samples annealed at a temperature of, say, $T_a = 120\,°\mathrm{C}$.

From Fig. 12 and eq. 53 we see that the basal surface tension $\sigma_e$ is almost independent of the average chain orientation in the amorphous phase (and in the interphase too). This is because *i)* the quantities $v_c \approx 140\ \mathrm{cm}^3/\mathrm{mol}$, $H_f \approx 2.5\ \mathrm{kJ/mol}$ and $T_m^\infty \simeq 285\,°\mathrm{C}$ are parameters characterizing the *crystalline phase*, and *ii)* the data of un-oriented PET follow the Gibbs-Thomson line fitting the oriented samples' data. (In fact, $\sigma_e$ should depend on the change in the average chain orientation when crossing the interphase from the crystalline to the amorphous region. Since however the interfacial thickness depends linearly on the density-density correlation length, the oriented PET is characterized by thicker basal interphases than the un-oriented PET. Thus, mean chain orientation change through the interphase, and thickness of the latter, seem to almost balance to yield a practically constant $\sigma_e$. This at least as it appears from the Gibbs-Thomson plot of Fig. 12. This conjecture, however, is anopen problem as yet.)

On the other hand, eq. 51 allows for the estimate of the excess chemical potential $\delta\mu = \Delta\mu' - \Delta\mu$ (see eq. 85) from the difference $\delta l_0$ between the lamellar thicknesses of the samples, i.e., $l_0 \simeq 2.7$ nm for the un-oriented sample and $l_0 - \delta l_0 \simeq 2$ nm for the oriented sample:

$$|\Delta\mu + \delta\mu| = \frac{2\sigma_e v_c}{l_0}\left(\frac{l_0}{l_0 - \delta l_0}\right) . \tag{90}$$

From the Gibbs-Thomson plot and from eq. 90, a straightforward calculation yields $|\Delta\mu| \approx 0.74\ \mathrm{kJ/mol}$ at $T = 120°\mathrm{C}$, and a value $|\delta\mu| \approx 0.26\ \mathrm{kJ/mol}$ to justify $\delta l_0 \approx 0.7$ nm. By inserting these values into eq. 85, assuming $\xi \approx 1$ nm (i.e. of order the monomer length), and considering that $\kappa = 4$, we find a cutoff wave vector of $k_c = 2.9 \times 10^7\ \mathrm{cm}^{-1}$. The corresponding (parallel) wavelength is

$2\,\pi/k_c \approx 2$ nm, i.e. twice as much the *ordinary* correlation length, as it should be (see eq. 87 and the comments just following it).

Of course there is a number of simplifying assumptions in the whole calculation developed above, and the agreement with the experiment should be considered satisfactory just because our estimates are not orders of magnitudes far from the data. In conclusion, we may take our reasoning as a possible basis upon which more detailed calculations (e.g. simulations) can be performed.

### 2.5.6 Appendix: Spinodal decomposition

We can illustrate the difference between nucleation and spinodal decomposition by resorting to the ideas of Ginzburg and Landau [6].

We describe phase transitions through an expansion of a phenomenological free energy $F$ in terms of a suitable *order parameter* $\phi$. To make things simple, let us consider a system able to undergo a phase transition, and whose local properties can be described by a conserved order parameter $\phi$ which is a function of just one spatial coordinate, say it $x$. The overall free energy $F$ is a functional of $\phi(x)$ given by

$$F[\phi] = \int d\,x \left[ f(\phi) + c \left( \frac{d\phi}{dx} \right)^2 \right] , \tag{91}$$

where the square gradient accounts for the energy contribution associated to the formation of an interface. Suppose that the free energy density $f(\phi)$ has the profile depicted in Fig. 19, and that the disordered phase is characterized by a value of $\phi$ larger than in the ordered one.

At high temperatures $T_h$, above the melting point, $f$ has only one minimum, say at $\phi = \phi_0$, corresponding to the equilibrium melt. As the temperature decreases down to $T_l$, below the melting point, $f$ assumes a two-well shape and the system tends to phase-separate into two states characterized, say, by $\phi_+ > \phi_0$ and $\phi_- < \phi_0$. If in these conditions $(\partial^2 f/\partial\,\phi^2)_{\phi=\phi_0} < 0$, the undercooled liquid is

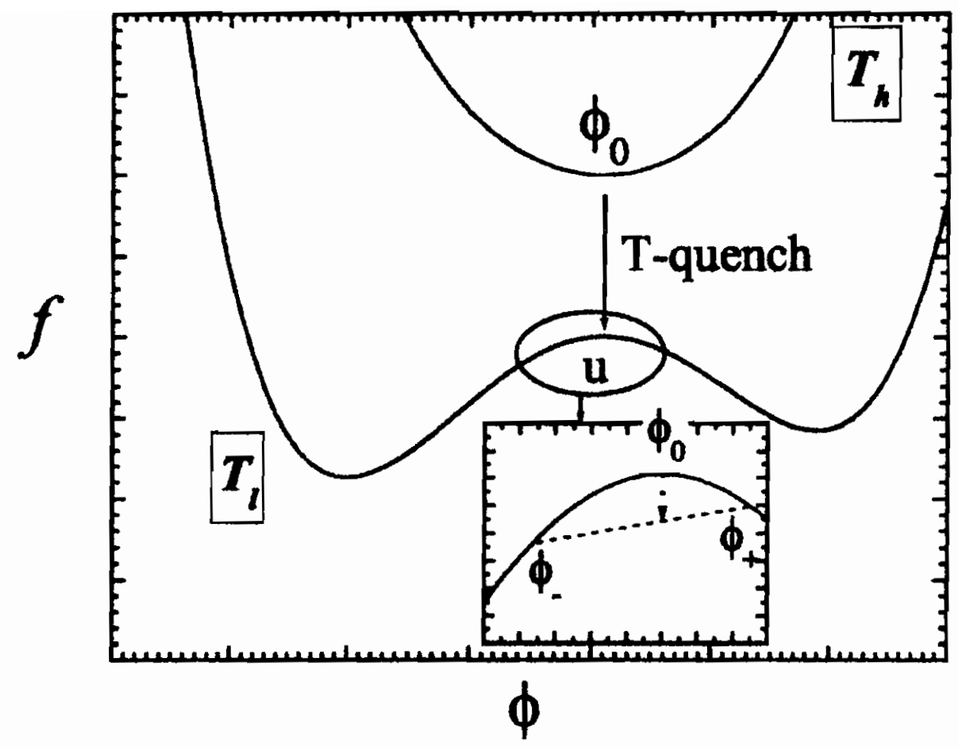



Figure 19: The free energy density $f$ as a function of the order parameter $\phi$ for high and low temperatures, $T_h$ and $T_l$ respectively. The temperature quenching drives the system form a condition of stable equilibrium at $T_h$, to a state of instability $\partial^2 f/\partial\phi_0^2 < 0$. From the Gibbs construction we see that phase separation into $\phi_+$ and $\phi_-$ regions is accompanied by an average decrease of $f$.

*unstable*: the transition towards the ordered state takes place via spinodal decomposition, a spontaneous process which, differently from nucleation, is not activated.

We consider the case where the order parameter is conserved (for instance it may describe a concentration of *defects*, i.e. entanglements or whatever else), so that its evolution must be compatible with the constraint $\int dx(\phi - \phi_0) = \text{const}$. We won't go into the details, and limit ourselves to show that a periodic modulation of the order parameter may lead to a decrease of the overall free energy, *provided* the gradient of $\phi$ is not too large, i.e. the period of the modulation is sufficiently long. So, let $\delta\phi \equiv \phi - \phi_0 \sim \cos(qx)$ be a small perturbation of the (unstable) equilibrium order parameter $\phi_0$; then, by expanding $f$ to lowest significant order in a Taylor

series about $\phi_0$, we find:

$$\Delta F[\phi] \equiv F[\phi] - F[\phi_0] \propto \left[ \left( \frac{\partial^2 f}{\partial \phi^2} \right)_{\phi=\phi_0} + cq^2 \right] . \qquad (92)$$

From this expression it is evident that $\Delta F[\phi] < 0$ provided

$$q < \sqrt{-\frac{2}{c} \left( \frac{\partial^2 f}{\partial \phi^2} \right)_{\phi=\phi_0}} . \qquad (93)$$

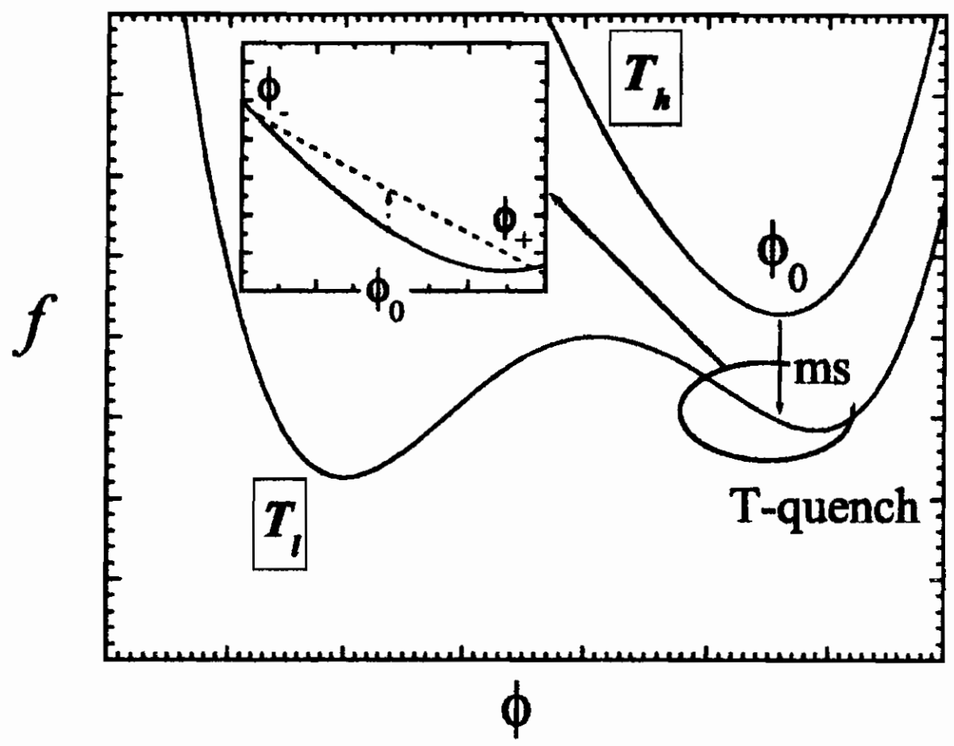



Figure 20: The free energy density $f$ as a function of the order parameter $\phi$ for high and low temperatures, $T_h$ and $T_l$ respectively. The temperature quenching drives the system form a condition of stable equilibrium at $T_h$, to a state of meta-stability $\partial^2 f/\partial \phi_0^2 > 0$. From the Gibbs construction we see that phase separation into $\phi_+$ and $\phi_-$ regions is accompanied by an average increase of $f$.

On the other hand, nucleation is the mechanism of phase separation when the $T$-quench drives the system into a region of meta-stability, where $(\partial^2 f/\partial \phi^2)_{\phi=\phi_0} > 0$. As shown in Fig.20, in this case the separation in regions characterized by different values of

the order parameter, $\phi_+ > \phi_0$ and $\phi_- < \phi_0$, is accompanied by an increase of the average free energy $f$, as the Gibbs construction points out.

# 3 Monohydric alcohols

## 3.1 Introduction

Monohydric alcohols are substances whose molecules are able to *associate* by means of a mechanism known as *hydrogen bonding.* This kind of binding interaction is responsible of the formation of supramolecular aggregates whose properties affects rather drastically their macroscopic behavior.

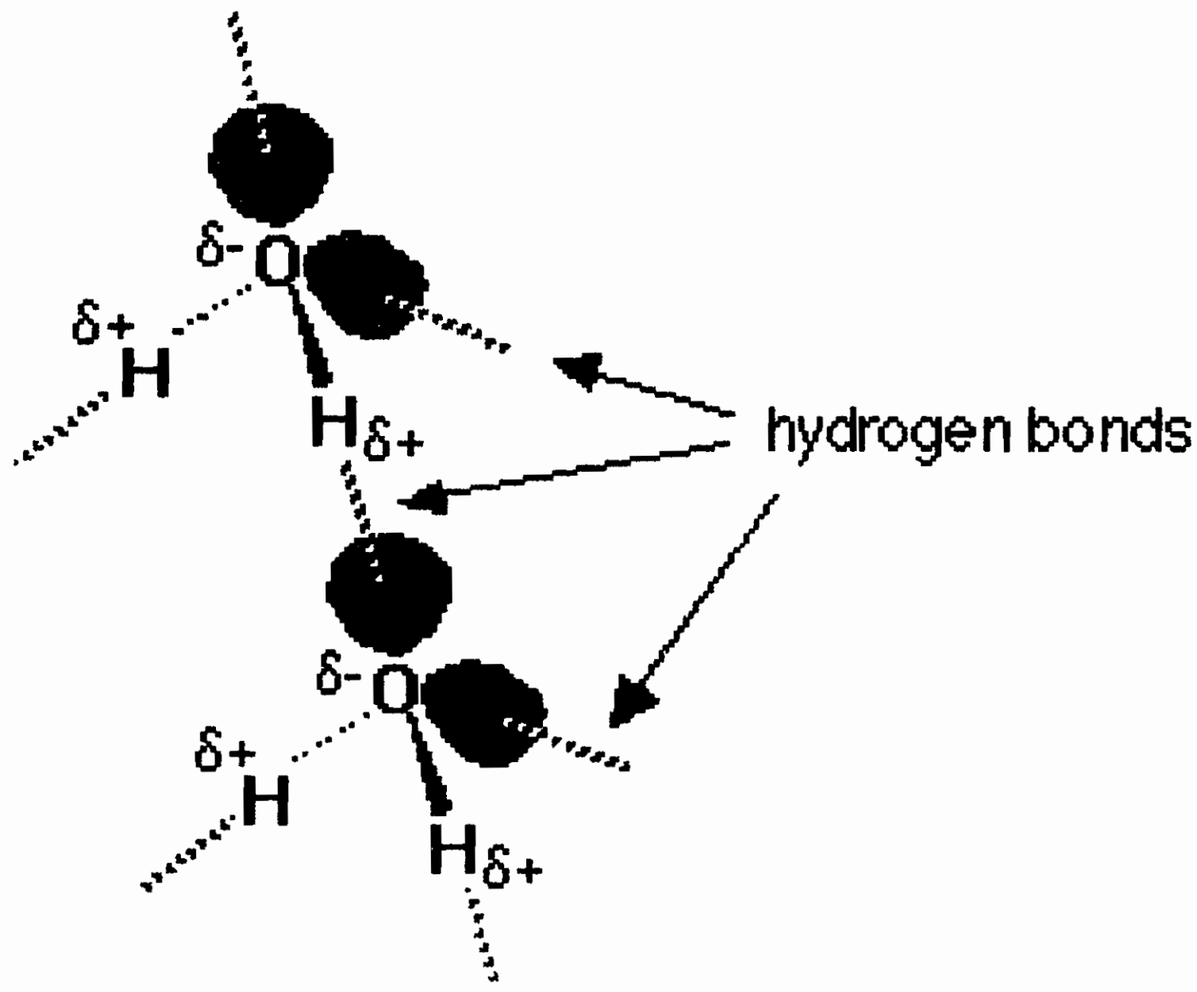



Figure 21: Water molecules bonded together via an hydrogen bond.

Water is also made of molecules undergoing mutual interaction via hydrogen bonding; for this reason this substance can be found in the liquid phase within so a wide temperature range at atmospheric pressure, and its boiling point is so high. Consider two water molecules coming close together. The $\delta+$ hydrogen is so strongly attracted to the lone pair that it is almost as if you were beginning to form a co-ordinate (dative covalent) bond. It doesn't go that far, but the attraction is significantly stronger than an ordinary dipole-dipole interaction. Hydrogen bonds have about a tenth of

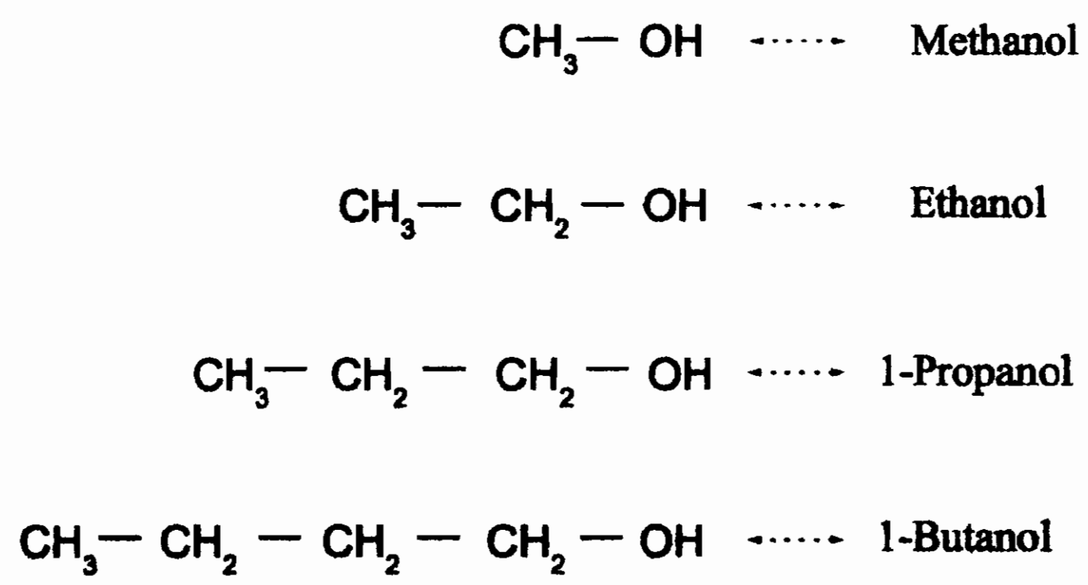



Figure 22: Linear monohydric alcohols with varying length of the alkyl tail.

the strength of an average covalent bond, and are being constantly broken and reformed in the liquid phase.

The interesting feature of monohydric alcohols is that in the liquid phase they form mainly chain-like aggregates, which offer the advantage of being rather easy to describe from the statistical mechanical point of view, even though their shape and dimension change restlessly.

In the present chapter we shall mainly consider the quasi-static dielectric properties of these liquids (and of their mixtures with a non polar solvent), and see how these properties can be interpreted in terms of the statistics of the aggregates.

Our starting point will be the experiment of course, but we shall see that the theoretical construction hinted by the need of a coherent interpretation of the data will go beyond the observations, in the spirit of the best traditions of the Physical Science.

## 3.2 Dielectrics and Calorimetry

### 3.2.1 The experiments

Calorimetry has been used to estimate the strength of the hydrogen bond in monohydric alcohols through the measurements of the enthalpy of mixing $\Delta H_{mix}$ in diluted alcohol-in-carbon tetrachloride mixtures [16]. Although carbon tetrachloride ($CCl_4$) is a non polar molecule (its structure is analogous to methane, but with chlorine put in place of the hydrogen atoms), it is a good solvent for alcohol molecules such as those shown in Fig. 22.

The procedure consists first in measuring the enthalpy $H$ of formation of an alcohol/$CCl_4$ mixture as a function of the alcohol mole fraction $x$, and then in calculating the difference

$$\Delta H_{mix} \equiv H - xH_a - (1-x)H_s\,, \tag{94}$$

where $H_a$ and $H_s$ are the enthalpies of the alcohol and the solvent respectively, and $x$ is the mole fraction of the alcohol. By extrapolating $\Delta H_{mix}(x)$ for vanishing $x$, the partial molar mixing enthalpy of the alcohol component can be obtained. Indeed, the effect of an increasing dilution is the rupture of a progressively increasing fraction of hydrogen bonds binding the alcohol molecules together, and the amount of energy needed to break these bonds is connected with the heat to be supplied in the calorimeter.

This procedure was followed for solutions of methanol, 1-propanol, 1-butanol and 1-octanol in $CCl_4$; the resulting $\Delta H_a^\infty$ was invariably found to fall within the interval 4-6 kcal/mol. Now, the final volume of each mixture did not differ more than 0.1 % from the sum of the volumes of the components before mixing, so that $\Delta H_a^\infty$ was in fact a measure of the binding *energy*. The independence of this result on the length of the alkyl tail suggested that the partial heat of mixing $\Delta H_a^\infty$ was in fact to be related to some process involving the hydroxyl group of the molecules: the only feature shared by the whole variety of alcohol species analyzed.

So, calorimetry first pointed out a sort of general behavior characterizing a moderately wide class of molecules, namely, the strength

of the hydrogen bonds they may form.

We now leave calorimetry and consider the quasi-static dielectric response of these liquid mixtures (quasi-static means that the dielectric response is not really measured in steady conditions, but rather at a frequency of 100 kHz; the period of the field, however, may be considered virtually infinite as compared to the relevant characteristic time scales of the molecular processes occurring in the temperature intervals we are interested in).

Figure 23 reports the relative dielectric constant $\varepsilon$ as a function of the alcohol volume fraction $\phi \equiv V_a/(V_a + V_s)$, being $V_a$ and $V_s$ the volumes of the alcohol and the solvent used to make the solution. The measurements have been carried out at room temperature in mixtures of methanol, ethanol, 1-propanol and 1-pentanol in $CCl_4$ [17]. We cannot reproduce here the general derivation, but for a mixture of a polar liquid and a non-polar solvent the dielectric constant $\varepsilon$ is related to the dielectric parameters of the components by [18]

$$\frac{9k_BT}{4\pi\varepsilon}\left\{\frac{\varepsilon-\varepsilon_{\infty,a}}{2\varepsilon+\varepsilon_{\infty,a}}\phi+\frac{\varepsilon-\varepsilon_{\infty,s}}{2\varepsilon+\varepsilon_{\infty,s}}(1-\phi)\right\}= \frac{N_A}{v_a}\left(\frac{\varepsilon_{\infty,a}+2}{2\varepsilon+\varepsilon_{\infty,a}}\right)^2\mu_{eff}^2\,\phi\,, \tag{95}$$

where $\varepsilon_{\infty,a}$ and $\varepsilon_{\infty,s}$ are the relative dielectric constants at optical frequencies of the alcohol and the solvent respectively, $N_A$ is the Avogadro number, $v_a$ is the molar volume of the alcohol and $\mu_{eff}$ is an effective electric dipole to be associated to each molecule of the *polar* component. This latter quantity is known to depend by a number of conditions, such as solvation and association. By *ideal* behavior we shall mean the case where the effective dipole $\mu_{eff}$ is a constant in the whole $\phi$-range.

The ideal behaviors predicted for the mixtures are reported in Fig. 23. In all cases the actual dielectric constant of each mixture is lower than ideal in the low-$\phi$ region. In the high-$\phi$ region the discrepancy between the actual and the ideal values of $\varepsilon$ depends

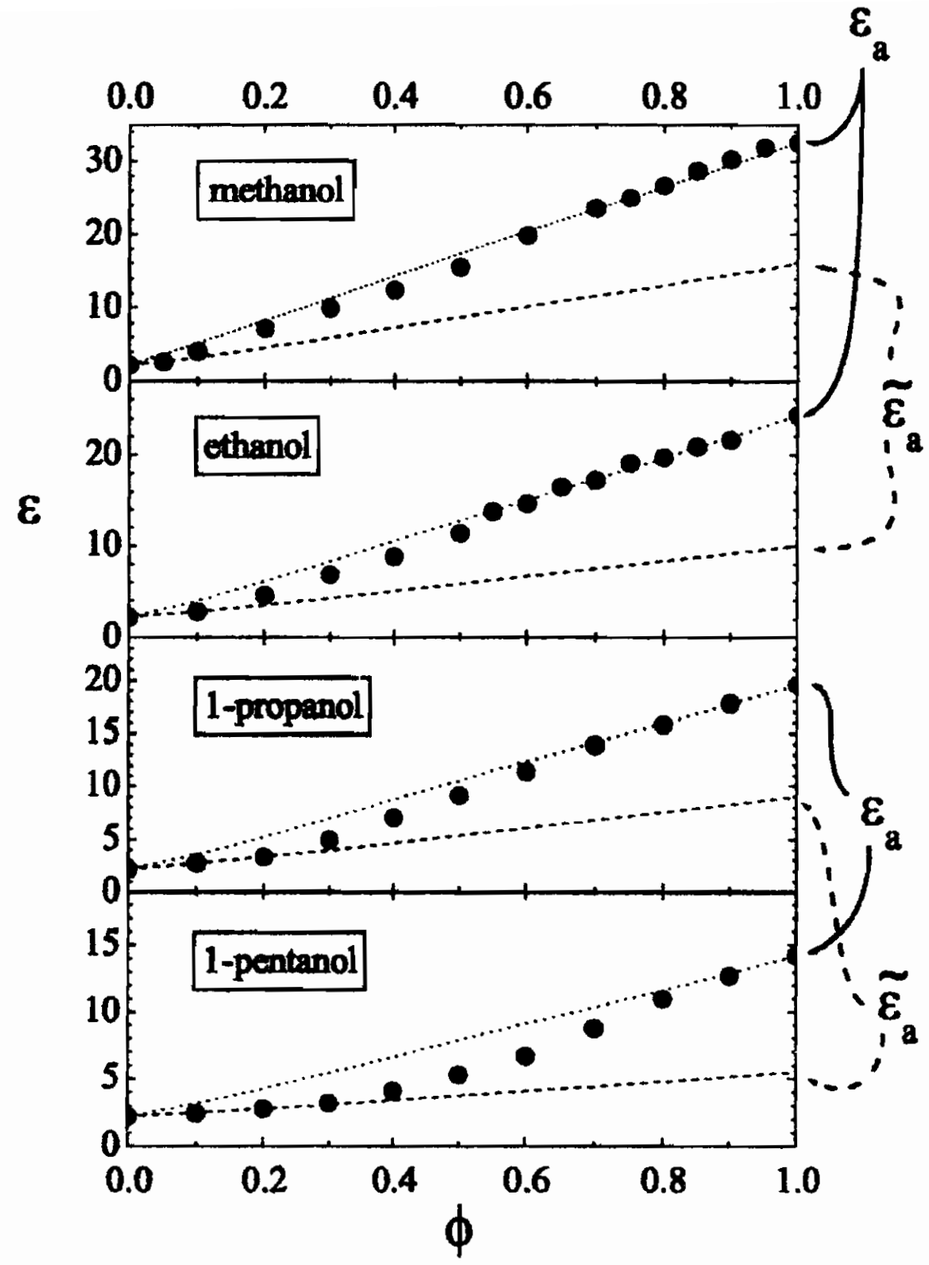



Figure 23: Room temperature relative dielectric constant in the alcohol/$CCl_4$ mixtures as a function of the alcohol volume fraction $\phi$. The lines are ideal behaviors describing a polar non-associating liquid in a non-polar solvent.

on the alcohol; we shall discuss in detail the case of methanol afterwards.

It is worth noticing that in all cases, if the low-$\phi$ data are fitted by ideal behaviors, then the extrapolations to $\phi = 1$ intercept the ordinates at values $\widetilde{\varepsilon}_a$ of the dielectric constant which are systematically lower by about one half with respect to the value $\varepsilon_a$ expected for the pure alcohol. The results are summarized in the following table [the values of $\mu_{eff}$ have been obtained by imposing $\varepsilon(\phi = 1) = \varepsilon_a$]:

If we look at both the data and the ideal extrapolations, we see

| | $v_a/N_A$ (Å$^3$) | $\varepsilon_a$ | $\tilde{\varepsilon}_a$ | $\varepsilon_{\infty,a}$ | $\mu_{eff}$ (D) |
|---|---|---|---|---|---|
| methanol | 67.3 | 32.5 | 17.2 | 1.766 | 2.98 |
| ethanol | 97 | 24.3 | 9.7 | 1.852 | 3 |
| 1-propanol | 124 | 19.6 | 9 | 1.918 | 2.98 |
| 1-pentanol | 180 | 14.2 | 5.5 | 1.988 | 2.96 |

Table 1: Molecular volume $v_a/N_A$, bulk dielectric constant $\varepsilon_a$, ideal extrapolation $\tilde{\varepsilon}_a$, optical dielectric constant and effective (apparent) molecular dipole $\mu_{eff}$ for methanol, ethanol, 1-propanol and 1-pentanol at room temperature. For $CCl_4$, $\tilde{\varepsilon}_s \simeq 2.2$. The dipole moment is expressed in debyes (1 debye= $10^{-18}$ e.s.u. $= 3.3356 \times 10^{-30}$ Cm). All considered alcohol species are characterized by almost the same dipole moment *in vacuum* $\mu_{vac} \simeq 1.7$ D.

that in the low-$\phi$ region, each alcohol behaves *as if* it had a reduced molecular dipole. On the other hand, it is known that the alcohols are able to associate so to form *cyclic* aggregates (or clusters). The latter have no net dipole moment, so the whole fraction of molecules involved in the formation of these closed aggregates represents in fact a further non polar component in the mixture.

It is hard not to imagine that the impressive regularity of the data reported in the table has to do with what observed in calorimetry, so we assume that this behavior is also a manifestation of the association mechanism.

Thus it is important to see how the number of cyclic clusters depends on the hydrogen bonding energy and on the alcohol volume fraction. Since both the calorimetric and dielectric behaviors appear to be independent of the length of the alkyl tail, we seek for a simplified model where the shape of the alcohol molecules is disregarded. Moreover, since our concern is the influence of the binding energy on the shape of the cluster, we shall not consider degrees of freedom other than configurational.

### 3.2.2 A simple lattice model

Given a volume $V$, let $\phi$ be the fraction of it which is occupied by the alcohol molecules; the rest is supposed to be filled by the non-polar solvent. Hydrogen bonds among alcohol molecules continuously break and reform, so that each cluster changes its shape and aggregation number restlessly.

Let $n_{k,i}$ be the number of clusters made of $k$ molecules ($k$-cluster) in either cyclic ($i = 1$) or open ($i = 2$) configuration. The problem is to find the equilibrium distribution of these clusters, given an energy $E_h$ characterizing the strength of the hydrogen bonding.

We do this by resorting to the method outlined at the end of Section 1.2.2: we calculate the number of distinguishable ways a given cluster distribution can be realized, and extremize it under the constraints that the average energy is a constant and that the total number of alcohol molecules is fixed (this is equivalent to find the extremum of the appropriate thermodynamic potential: the Lagrange multiplier associated to the constraint on the average energy will eventually assume the appropriate value).

Let $E_h$ be the energy needed to separate two mutually bonded alcohol molecules *in the presence of the solvent*, i.e. roughly the difference between the hydrogen binding, $E_{h0}$, and the solvation energies; then, a $k$-cluster must be assigned an overall formation energy

$$E_{k,i} = \begin{cases} -kE_h & i = 1 \ \text{(cyclic)} \\ & \qquad\qquad k > 2 \\ -(k-1)E_h & i = 2 \ \text{(open)} \end{cases} \tag{96}$$

$$E_k = -E_h \delta_{k,2} \qquad\qquad k \leq 2$$

i.e. simply $-E_h$ times the number of bonds. We exclude at the outset the possibility that dimers can form with two hydrogen bonds; these configurations are possible, but very unstable.

In order to estimate the number of distinguishable configurations which can be realized by a given distribution, we use the trick

to discretize positions and molecular orientations by means of a cubic lattice. This is a widespread used technique which simplifies the calculations considerably, and has been used already when we introduced the gaussian polymeric chain.

We imagine that the alcohol molecules can occupy only the vertices of the cells of a cubic lattice (coordination $z = 6$), and that their electric dipoles are allowed to be oriented only along the principal directions defined by the lattice. (We allow dimers to configure cyclic-like in the discretization of the lattice, still excluding however the possibility of more than one bond.) The volume of each cell equals the volume of a molecule.

Let $N_t$ denote the total number of cells in which the whole volume $V$ occupied by the mixture can be divided, and consider a chain made of $k$ molecules. Then, the number of ways the cluster can be accommodated in the lattice is

$$W_k = N_t z^k \,, \tag{97}$$

because there are $N_t$ possibilities to place the first molecule of the cluster, and the position of each other molecule depends on the dipole orientation of the preceding one. Note that if the $k$ distinguishable molecules were not connected, the much larger number $(N_t z)^k$ would have been found for $W_k$. Of course, both the open configurations and the cyclic ones are included in $W_k$; the latter, as we saw, are expected to be relevant in determining the the deviation from ideality of the dielectric response of our liquid mixtures, so it is important to find how many configurations out of $W_t$ correspond to closed chains.

A simple calculation yields the following expression for the fraction $P_k$ of cyclic configurations out of $W_k$ (see appendix):

$$P_k = \frac{k!}{z^k} \sum_{j,l=0}^{k/2} \left[ l!\, j!\, \left( \frac{k}{2} - j - l \right)! \right]^{-2} \;\; ; \;\; \begin{cases} j + l \leq \frac{k}{2} \\ \\ k \;\; \text{even} \end{cases} \tag{98}$$

Strictly speaking, cyclic clusters cannot form in a cubic lattice if $k$ is odd; however, the constraint imposed by the lattice geometry

is fictitious, since it only serves as a convenient way to discretize configurations and estimate entropy differences easily. Thus we shall use eq. 98 for $k$ even, and use a suitable interpolation for $k$ odd. Figure 24 shows how fast $P_k$ decays with increasing $k$.

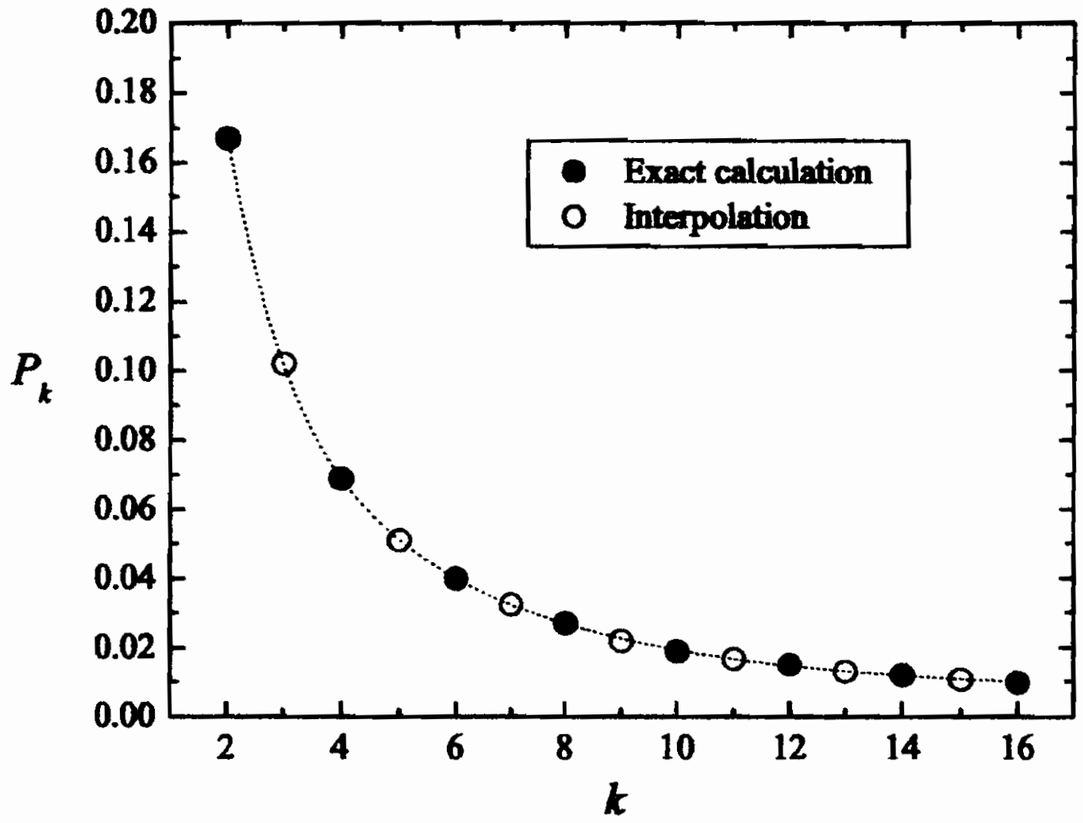



Figure 24: Fraction $P_k$ of cyclic configurations out of $W_k$. The full symbols are obtained from eq. 98, the open ones have been obtained from an interpolating sum of two exponential decays.

We can now calculate the number $W_{k,i}$ of *distinguishable* configurations that a $k$-cluster may take:

$$W_{k,i} \equiv \frac{W_k}{2} \times \begin{cases} P_k(k-\delta_{k,2})^{-1} & i=1 \ \text{(cyclic)} \\ 1-P_k & i=2 \ \text{(open)} \end{cases} \tag{99}$$

where the factor $1/2$ accounts for the two possible molecular ordering along the same path, and $(k-\delta_{k,2})^{-1}$ is to get rid of the redundancy arising from the equivalence of choosing any of the molecules as the first one in a cyclic configuration.

Apart of irrelevant factors, the total number of distinguishable configurations that the whole cluster distribution $\{n_{k,i}\}$ may as-

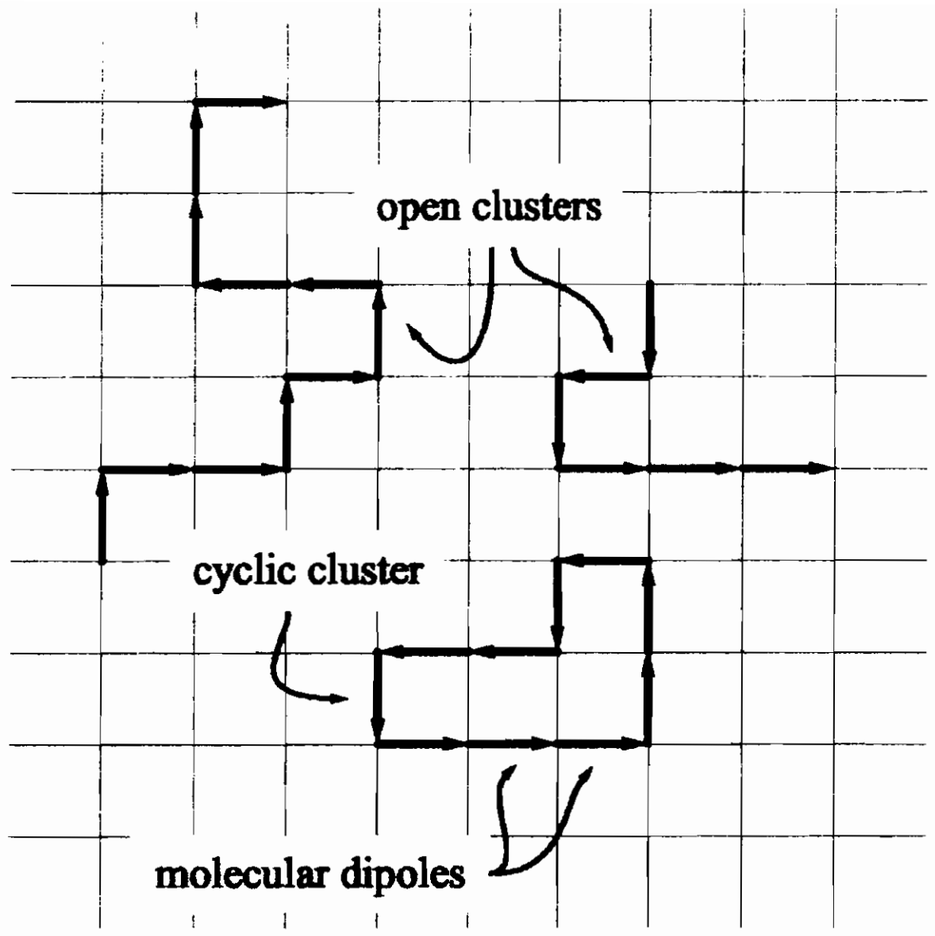



Figure 25: Open and cyclic clusters in a two-dimensional lattice.

sume is given by

$$W_t = \prod_{k,i} \frac{(W_{k,i})^{n_{k,i}}}{n_{k,i}!} \,. \tag{100}$$

The entropy of the system is then $S = k_B \ln W_t$, and the equilibrium condition requires that $S$ be a maximum under the following constraints:

1. the total number $N$ of alcohol molecules is fixed:

$$n_1 = N - \sum_{k=2,\,N;\,i=1,\,2} k\; n_{k,i} \tag{101}$$

($n_1$ is the number of non-bonded molecules, or 1-clusters; the combination $k = 2$ and $i = 1$ is excluded from the summation).

2. Both the total volume $V$ and the alcohol volume fraction $\phi$ are fixed, so that all terms in the relationship

$$N = \phi N_t \tag{102}$$

are fixed.

3. The *average* overall energy

$$U = -E_h \sum_{k=2,\,N;\,i=1,\,2} n_{k,i}\{\delta_{k,2} + (k - \delta_{i,2})(1 - \delta_{k,2})\} \quad (103)$$

must be a constant.

We obtain the equilibrium distribution by simply finding the maximum of $S + \alpha U$ with respect to the arbitrary variations of all $n_{k,i}$ with $k \geq 2$ (remind eq. 101), being $\alpha$ a Lagrange multiplier which eventually must be assigned the value $-1/T$. The resulting distribution is

$$n_{k,i} = n_1^k \frac{W_{k,i}}{W_1^k} \exp\left\{\frac{E_h}{k_B T}\left[\delta_{k,2} + (k - \delta_{i,2})(1 - \delta_{k,2})\right]\right\} . \quad (104)$$

This is just a "formal" solution of our problem; to find the actual one we must substitute eq. 104 back into eq. 101 and solve a polynomial of degree $N$ in the unknown $n_1$. Of course this can only be done numerically; however, we don't actually need to really find the exact solution. It is enough to truncate the sum in eq. 101 up to a value $\overline{N}$ which is significantly larger than the resulting average aggregation number

$$\langle n \rangle \equiv \frac{n_1 + \sum_{k=2,\,\overline{N};\,i=1,\,2} k\; n_{k,i}}{n_1 + \sum_{k=2,\,\overline{N};\,i=1,\,2} n_{k,i}} . \quad (105)$$

In practice, one increases $\overline{N}$ progressively until a rather stable value of $\langle n \rangle$ is found.

Although eq. 104 is only a formal expression, it is nevertheless indicative of a very important fact, namely, that $n_{k,i}$ does not depend explicitly on $z$, i.e. on the discretization imposed by the lattice. This is because the factors $z^k$ in $W_k$ and in $W_1^k$ cancel out. There is in fact an implicit dependence on $z$ through $P_k$ but it is effective only for small values of $k$, and also in this case, the agreement of the results with numerical simulations is excellent also in these conditions [19].

There are a few more things which deserve being pointed out. The first one is that it is implicit in our calculations that there is no correlation between the direction of a dipole whatsoever and those of the preceding or of the subsequent one along the chain. In other words there is no energy assigned to the angle that two successive dipoles in a chain may form, thus all possible relative orientations are equiprobable. This has been implicitly assumed while assigning equal probability and equal energy to all possible configurations $W_t$ (for a fixed $\{n_{k,i}\}$) when writing the expression of the entropy. This is of course an approximation, but it does its job very well as we shall see; from now on it will be referred to as the "freely jointed chain approximation".

Another point is worth be noticed, namely, that in writing down the number of distinguishable configurations, we have in fact treated our clusters as "ideal" chains, allowing for multiple occupation of any lattice site whatsoever. This is in fact not a problem, because the constraint that both $V$ and $\phi$, *and not their averages*, are fixed, prevents multiple occupation at the time when the entropy is extremized.

Once the solution of our problem has been found, we have the possibility to calculate a number of quantities of relevance as functions of just two parameters, namely $\phi$ and the ratio $r \equiv E_h/k_BT$. Figure 26 reports the average aggregation number $\langle n \rangle$, as given by eq. 105, for varying $\phi$ and for three different values of $r$. The average numbers $\langle n_c \rangle$ and $\langle n_o \rangle$ of the cyclic and open clusters can be found from the same equation by restricting the sums to either $i = 1$ or $i = 2$ respectively (in the case $i = 1$ the term $n_1$ is ruled out, of course).

Another very important quantity which can be calculated rather easily is the fraction of molecules involved in the formation of cyclic clusters:

$$\eta \equiv \frac{\sum_{k=3,\overline{N}} k\ n_{k,1}}{n_1 + \sum_{k=2,\overline{N};i=1,2} k\ n_{k,i}} \tag{106}$$

Figure 27 reports $\eta$ as a function of $\phi$ for varying $r$. Note the tendency of $\eta$ to increase up to a maximum, and then to decrease

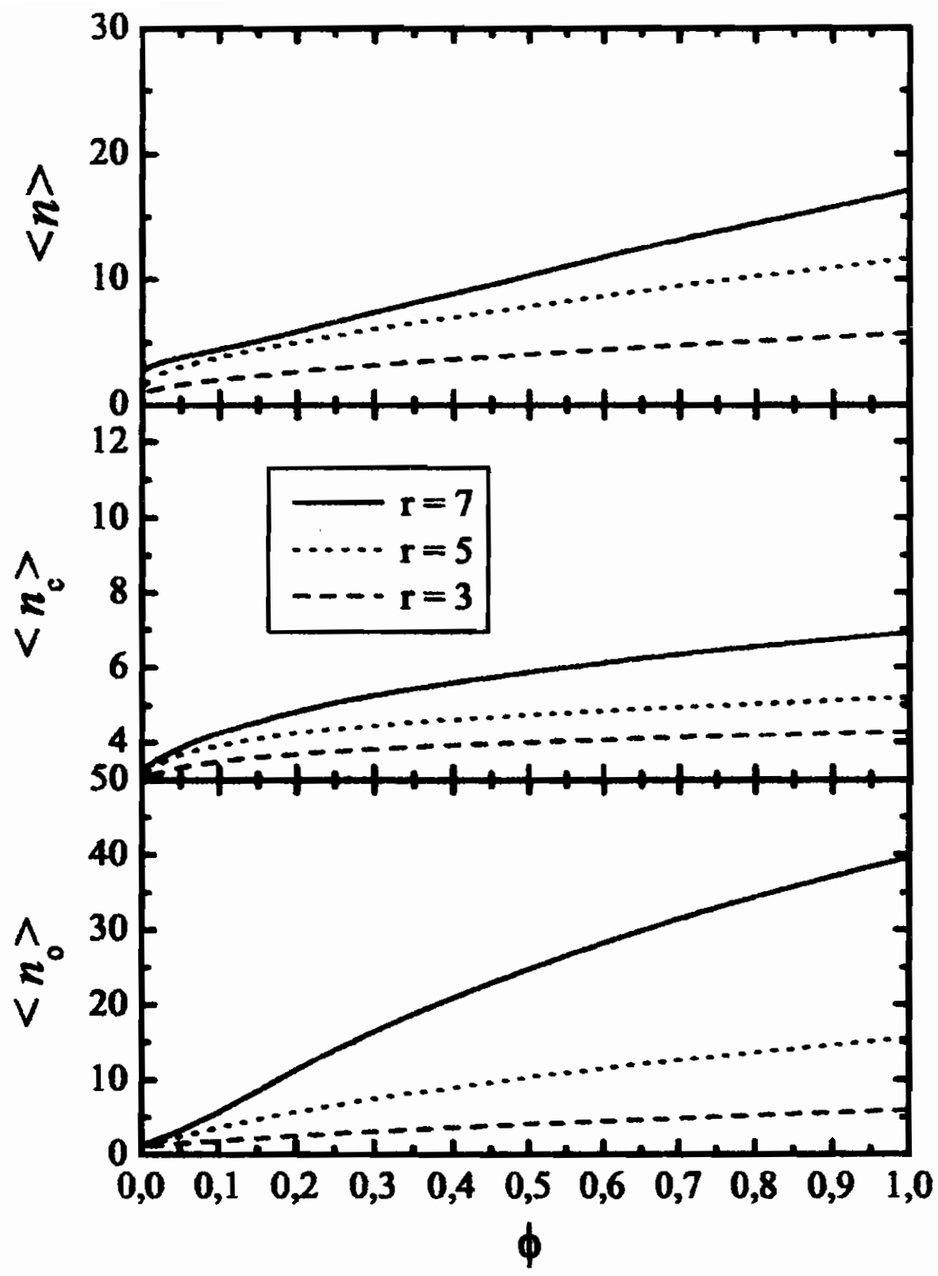



Figure 26: Overall average aggregation number $\langle n \rangle$, eq. 105, as a function of the alcohol volume fraction $\phi$, for three different values of $r \equiv E_h/k_BT$ (see main text). The sum has been truncated at $\overline{N} = 1000$. In the other panels the average aggregation numbers $\langle n_c \rangle$ and $\langle n_o \rangle$ of the cyclic and open clusters respectively are reported for the same values of the parameters.

towards zero for decreasing $\phi$, as opposed to the behavior of the average aggregation number shown in Fig. 27. The physical reason for this effect is that the probability of a $k$-cluster to be in a closed or open configuration, results from a competition between the energy decrease ($E_h$) due to the mutual binding of the chain ends of the cluster, and the corresponding configurational entropy loss. The latter overwhelms the former for large $k$, but the two become comparable as the clusters decrease their length, until at extremely

low $\phi$ any binding is accompanied by too a large entropy decrease.

It is important to note that in the low-$\phi$ region, it is possible that the majority of the alcohol molecules is involved in cyclic clusters. It is also worth noticing that a non-negligible population of cyclic clusters is present even in the case that the alcohol is pure, and it increases significantly with decreasing the temperature. We shall see afterwards that this is an important point to consider for a more detailed analysis of the dielectric behavior of the alcohols.

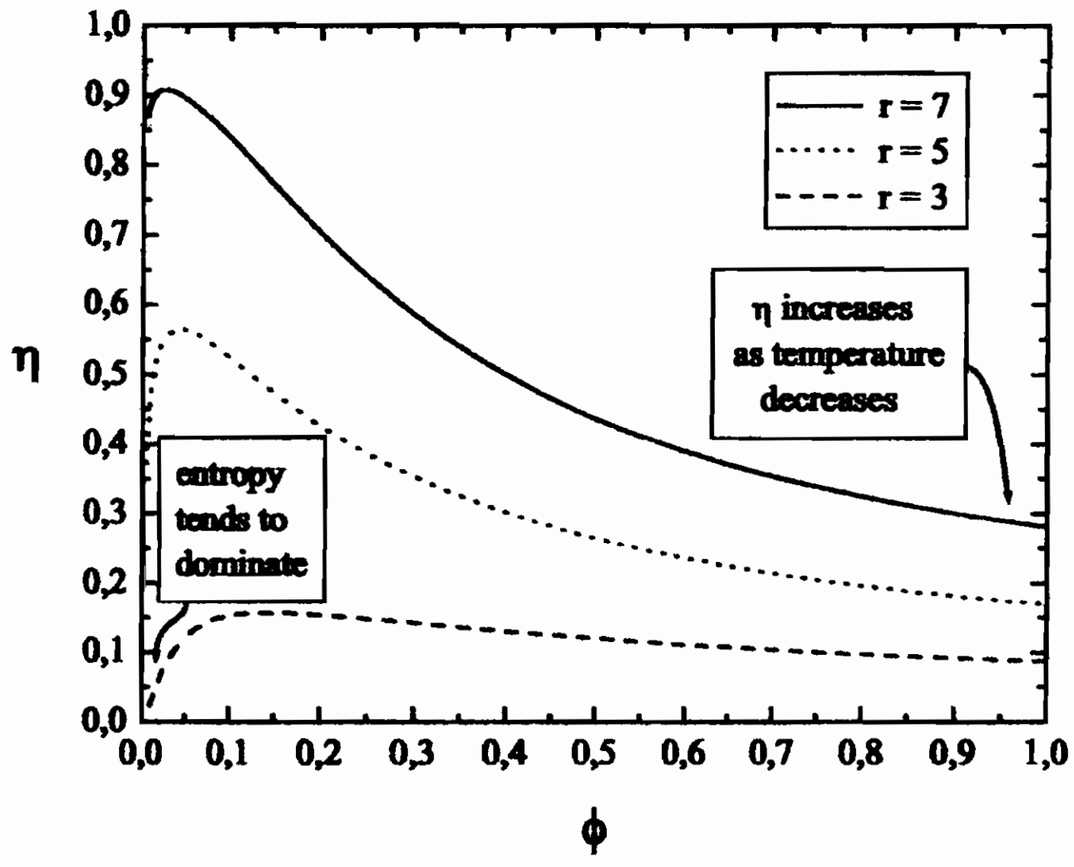



Figure 27: The fraction of alcohol molecules associated to form cyclic clusters is a function of $\phi$ and of the binding energy normalized to $k_BT$.

### 3.2.3 Comparison with the experiment

We see from Fig. 27 that the fraction of molecules forming cyclic clusters can really be large. If this is the main mechanism responsible of the systematic deviations from ideality of the actual dielectric constant, then we are facing with two main issues:

1. Eq. 95 is wrong, because the volume fraction of the component which "really" contributes as polar is not $\phi$ anymore (i.e. the

nominal alcohol volume fraction), but roughly $\phi(1-\eta)$, so we must make some correction.

2. Since the actual values of $\eta$ depend on the ratio $E_h/k_BT$, we must check whether the energies implied by the matching of the corrected equation with the data are close enough to the results of calorimetry.

About the first point we just put $\phi(1-\eta)$ in place of $\phi$ in the r.h.s. of eq. 95; moreover, since the optical dielectric constants $\varepsilon_{\infty,a}$ of the alcohols are very close to that of $CCl_4$, we just consider them to be the same. So we are led to the equation

$$\frac{(\varepsilon-\varepsilon_{\infty,s})(2\varepsilon+\varepsilon_{\infty,s})}{\varepsilon(\varepsilon_{\infty,s}+2)^2}=\frac{4\pi N_A}{9k_BTv_a}\mu_{eff}^2\phi(1-\eta)\,, \tag{107}$$

where we chose to put $\varepsilon_{\infty,s}$ in the left hand side in order to match the dielectric constant of $CCl_4$ in the limit $\phi\to 0$.
We now have to answer the following question:
"Which value of $E_h/k_BT$, *within our model*, would be able to justify the apparent $\tilde{\varepsilon}_a$ values of Table 1?"
It is of no use to make rigorously correct calculations in this context, so we assume just for now, that at $\phi=1$ there are no cyclic clusters, i.e. that all alcohol molecules behave as polar (accounting for the presence of the non polar component inherent to the association mechanism would certainly be more correct, but since we are seeking for a loose agreement presently, we leave this complication out of our thoughts). In this case we can estimate

$$1-\eta\simeq\frac{\tilde{\varepsilon}_a}{\varepsilon_a}\,, \tag{108}$$

and from Table 1 we find $0.39\lesssim(1-\eta)\lesssim 0.53$, that is, $0.47\lesssim\eta\lesssim 0.61$. About the actual $\phi$ at which $\eta$ is expected to be in this range, we take the highest possible value *common to all considered alcohol species*, for which the data follow the ideal behavior connecting $\varepsilon_{\infty,s}$ to the respective $\tilde{\varepsilon}_a$. The maximum common value is that of

methanol, i.e. $\phi \simeq 0.1$. Now, from Fig. 27 we see that for this value of $\phi$, $r$ must be about 5 for $\eta$ to fall within the correct interval. Since at room temperature $k_B T$ corresponds to 0.6 kcal/mol, we find $E_h \approx 3$ kcal/mol (by accounting for the presence of cyclic clusters also at $\phi = 1$ we would have found a slightly higher energy value). This estimate must be considered to really be very close to the results of the calorimetric investigations.

In conclusion, we found that the model developed above represents a means to point out the consistency between calorimetric and dielectric observations. The model is very simple, but it seems to catch the relevant features of association with regards to the phenomenologies considered so far.

Most important, all this story points out the central role of the clusters' *configurational statistics* in determining the dielectric response of the alcohol species which we considered above. Of course, we analyzed just the low-$\phi$ region, but our results hint at the possibility that the role of the configuration of clusters might be important also in pure alcohols, as in fact we shall see.

### 3.2.4 Appendix

In this subsection we outline the calculation procedure to obtain the analytical expression of $P_k$. As explained in the text, this can only be done for $k$ even, since we have chosen to discretize in a cubic lattice.

Let $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$ and $\hat{\mathbf{z}}$ be the principal unit vectors of the lattice, and let $j_\pm$, $l_\pm$ and $m_\pm$ be the number of chain dipoles in a $k$-cluster oriented along $\pm\hat{\mathbf{x}}$, $\pm\hat{\mathbf{y}}$ and $\pm\hat{\mathbf{z}}$ respectively. We count the succession of the dipoles following the natural order and, if the $k$-cluster is closed, we chose the first dipole at will.

If the $k$-cluster has a closed configuration, then the number of dipoles along an arbitrary direction whatsoever must be equal to

the number of dipoles oriented in the opposite sense:

$$\begin{cases} j_+ = j_- \equiv j \\ \\ l_+ = l_- \equiv l \\ \\ m_+ = m_- = \frac{k}{2} - j - l \end{cases} \tag{109}$$

In order to find the number of possible configurations satisfying the above condition, we first calculate the total number of configurations with assigned $j_+$, $j_-$, ... $m_-$. To do this, we take our ordered sequence of dipoles and think about the direction of the $n$-th dipole as one among $z$ possible outcomes of a completely random variable. We are thus led to consider a binomial distribution of events; so, the number $W_{k;j_+}$ of configurations with assigned $j_+$ out of $W_k$ is given by

$$W_{k;j_+} = \frac{k!}{j_+!\,(k-j_+)!}\left[p_z^{j_+}(1-p_z)^{k-j_+}\right] W_k\,, \tag{110}$$

where $p_z$ is the *a priori* probability for the orientation of a dipole to be along the positive $\hat{\mathbf{x}}$ direction. Since there are $z$ equally probable available orientations, $p_z = 1/z$ (note that "equally probable orientations" means freely jointed chain, without steric hindrance).

Among the number $W_{k;j_+}$ of configurations just obtained, those with given $j_-$ can be enumerated similarly:

$$W_{k;j_+;j_-} = \frac{(k-j_+)!}{j_-!\,(k-j_+-j_-)!}\left[p_{z-1}^{j_-}(1-p_{z-1})^{k-j_+-j_-}\right] W_{k;j_+}\,, \tag{111}$$

where $p_{z-1} = 1/(z-1)$ is the *a priori* probability that the dipole is directed along $-\hat{\mathbf{x}}$ (remind the positive $\hat{\mathbf{x}}$ direction is not left available anymore after the calculation of $W_{k;j_+}$).

The procedure goes along the same lines to obtain at the end the number of configurations with assigned $j_+$, $j_-$, ... $m_-$; then, by applying the conditions eq. 109, we find:

$$[W_{k;j,l}]_{cycl.} = \frac{k!}{z^k}\frac{W_k}{\left[j!\,l!\,\left(\frac{k}{2}-j-l\right)!\right]^2} \tag{112}$$

The overall fraction of cyclic configurations out of $W_k$ is then

$$P_k = \frac{1}{W_k} \sum_{j,l=0}^{k/2} W_{k;j,l} \; ; \quad \begin{cases} j + l \leq \frac{k}{2} \\ \\ k \text{ even} \end{cases} \tag{113}$$

which is eq. 98.

## 3.3 Pure alcohols and chain stiffness

### 3.3.1 Introduction

In the freely jointed chain model the relative orientations of different dipoles are completely uncorrelated, even if the dipoles are mutually bonded. This oversimplification has its advantages in that it allows for a rapid estimate of the cluster distribution, without the need to resort to more complicated numerical procedures (such as simulations). However, we know already that an important role in dielectrics is played by dipole-dipole correlations, as we shall see shortly. Moreover, may be that we could find something interesting if we improve the model, where "interesting" means something which goes beyond the mere reduction of the error in the estimates from data fitting.

If we apply an electric field $\mathbf{E}$ to a homogeneous isotropic liquid, we shall induce an average polarization $p\,|\mathbf{E}|$. The relationship between $p$ and the dielectric constant is

$$\varepsilon - 1 = 4\pi \frac{N_A}{v_{liq}} p \,, \tag{114}$$

where $v_{liq}$ is the molar volume of the liquid.

In polar liquids, the permanent molecular dipole $\boldsymbol{\mu}$ tends to align along the direction $\mathbf{e} \equiv \mathbf{E}/|\mathbf{E}|$ of the applied electric field, and this mechanism contributes significantly to $p$ . If $\mathbf{M}$ is the total dipole moment induced on a specimen in the (homogeneous) field generated by a *fixed charge distribution*, and if $\mathbf{E}_0$ is the field

generated by the same charge distribution in the absence of the specimen, then the average component of $\boldsymbol{\mu}$ along $\mathbf{e}$ is:

$$\langle \boldsymbol{\mu} \cdot \mathbf{e} \rangle = \frac{\mathrm{Tr}\left\{ \boldsymbol{\mu} \cdot \mathbf{e} \exp\left[ -(U_{int} - \mathbf{M} \cdot \mathbf{E}_0)/k_B T \right] \right\}}{\mathrm{Tr}\left\{ \exp\left[ -(U_{int} - \mathbf{M} \cdot \mathbf{E}_0)/k_B T \right] \right\}} , \tag{115}$$

where $U_{int}$ is a many-body energy, comprising all molecular interactions, to be added to the contribution $-\mathbf{M} \cdot \mathbf{E}_0$ determined by the field; $\mathrm{Tr}\{\cdot\cdot\}$ stands for integration over the coordinates of the molecules in the specimen [20].

We can expand eq. 115 to lowest order in the field intensity to get:

$$\langle \boldsymbol{\mu} \cdot \mathbf{e} \rangle \simeq \frac{1}{k_B T} \langle \boldsymbol{\mu} \cdot \mathbf{M} \rangle_0 \mathbf{E}_0 , \tag{116}$$

where $\langle \cdot\cdot \rangle_0$ stands for the statistical average *in the absence* of the field, $\boldsymbol{\mu}$ is now the dipole moment of a given *fixed* molecule and $\mathbf{M}$ is the moment determined by $\boldsymbol{\mu}$ in the whole specimen through the intermolecular forces associated to $U_{int}$ [20].

Thus we see that in considering the component of $\boldsymbol{\mu}$ along the field, we have to account also for the interaction of each molecule with the others, and it is evident that this point should be of great relevance in the case of associating systems like alcohols.

Again, we cannot develop the theory here, but if the mechanism of correlations is accounted for properly, then the dielectric constant for a pure liquid satisfies the following equation

$$\frac{(\varepsilon - \varepsilon_{\infty,a})(2\varepsilon + \varepsilon_{\infty,a})}{\varepsilon(\varepsilon_{\infty,a} + 2)^2} = \frac{4\pi N_A}{9 k_B T v_a} \mu_0^2 g_{\mathrm{K}} (1 - \eta) , \tag{117}$$

where the factor $(1 - \eta)$ in the r.h.s. has been included to account for cyclic clusters (a term which can be found in a very limited number of papers at present [21]), $\mu_0$ is the molecular dipole *in the liquid phase*, and the Kirkwood correlation factor is defined by:

$$g_{\mathrm{K}} \equiv \frac{\langle \boldsymbol{\mu}_0 \cdot \mathbf{M} \rangle_0}{|\boldsymbol{\mu}_0|^2} . \tag{118}$$

It is our aim to derive $g_{\mathrm{K}}$ through a suitable improvement of our lattice model.

### 3.3.2 Chain stiffness

Dipole-dipole correlations are the result of complicated interaction mechanisms. The first we can mention is the short-range interaction between mutually bonded dipoles; but this is not the only one, because we don't have to forget the whole of the electrostatic interactions, which are very complicated to describe at molecular level.

We take a shortcut and make the following assumptions:

1. Dipoles belonging to different clusters are uncorrelated.

2. The probability distribution of the orientational correlations of adjoining dipoles (along the chain) are described by a *local* energy parameter associated to the hydrogen bond connecting the dipoles themselves.

About point (1), remind that $g_{\mathrm{K}}$ is a functions of averages calculated in the absence of the external field; in this situation we admit as reasonable to assume that the clusters are randomly conformed. Point (2) translates the idea of a mean field phenomenological approach which we shall briefly comment afterwards.

If we were able to recast the model accounting for the energies associated to the relative orientations of mutually bonded dipoles, we would have obtained at least a formal solution for the cluster distribution in a closed form. Instead we are not; so we consider a single cluster made of a suitably large number of dipoles, and assume it to be a stable subsystem for which statistical thermodynamics can be applied. Of course, $g_{\mathrm{K}} = 0$ for the dipoles of a cyclic cluster (see eq. 117), so we consider an open $k$-cluster.

As shown in Fig. 28, the space discretization imposed by the lattice, limits the mutual orientations (*bond states*) of two bonded dipoles to just three possible kinds, labelled with "$s$", "$b$" and "$f$" (to mean *straight*, *bend* and *fold* configurations respectively). States "$s$" and "$f$" have multiplicity *one* for a given orientation of one of the dipoles; on the other hand, states of the type "$b$" have multiplicity *four* because fixing the orientation of one of the dipoles

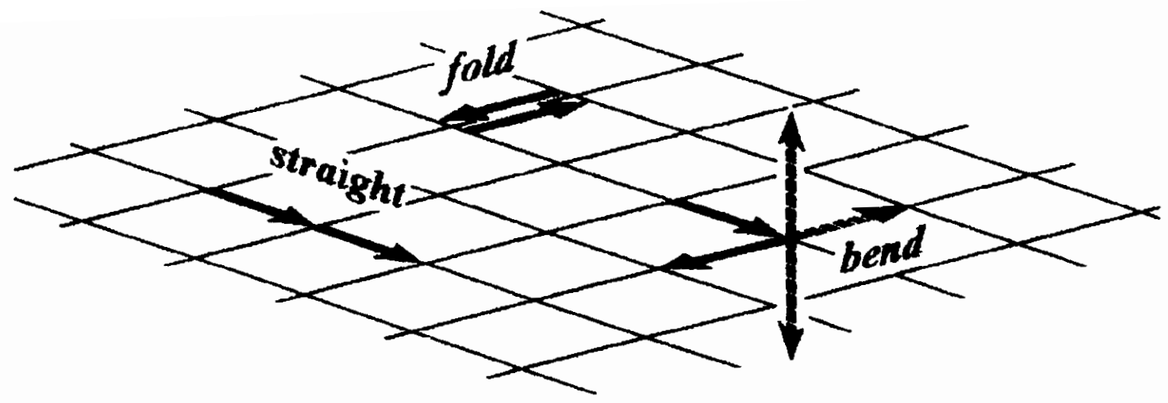



Figure 28: The three possible bond states in a cubic lattice.

leaves the possibility to the other to point towards the remaining four directions still available.

We take the energy $E_s$ associated to the "$s$" bond state as the reference and put $E_s \equiv 0$; $E_b$ and $E_f$ are the energies associated to the "$b$" and "$f$" states respectively. Now, point (2) above means that the bonds of a $k$-cluster constitute a set of independent subsystems of the $k$-cluster itself; in other words, there is no steric hindrance, the state of each bond is independent of the state of the others. We shall refer to this condition as to the *Markov assumption*. So, we are in the condition to write down the (canonical) partition function for a single bond:

$$Z_c = \sum_{\nu} [1 + (z-3)\delta_{\nu,b}] e^{-E_\nu / k_B T} \ , \quad \nu \equiv s,\, b,\, f \ , \tag{119}$$

in terms of which the bond state probabilities are

$$p_\nu = Z_c^{-1} e^{-E_\nu / k_B T} \ . \tag{120}$$

We can now calculate the orientational probability distribution of a chain dipole whatsoever once the orientation of another one (not necessarily adjoining) in the same chain is given. Let the labels $i = 1, 2, \ldots 6$ be assigned to the possible orientations $+\hat{\mathbf{x}}$, $-\hat{\mathbf{x}}$, ... , $-\hat{\mathbf{z}}$ respectively, allowed by the lattice. Then, let $\mathbf{q}$ be a probability 6-vector associated to a dipole, whose components

$q_i$ are the probabilities that the dipole itself is oriented along $+\hat{\mathbf{x}}$, $-\hat{\mathbf{x}}$, ..., $-\hat{\mathbf{z}}$. Now, assume that the $j$-th dipole of the $k$-cluster is oriented along, say, the $+\hat{\mathbf{x}}$ direction, which means that its 6-vector is $\mathbf{q}(j) = (1, 0, \ldots, 0)$. Then, the orientation probabilities of the $(j+n)$-th dipole of the same chain are the components of

$$\mathbf{q}(j+n) = \mathbf{A}^n \cdot \mathbf{q}(j)\,, \tag{121}$$

resulting from the $n$-fold iteration of the transformation defined by the matrix

$$\mathbf{A} = \begin{pmatrix} \mathbf{y} & \mathbf{w} & \mathbf{w} \\ \mathbf{w} & \mathbf{y} & \mathbf{w} \\ \mathbf{w} & \mathbf{w} & \mathbf{y} \end{pmatrix} \tag{122}$$

whose elements are

$$\mathbf{y} \equiv \begin{pmatrix} p_s & p_f \\ p_f & p_s \end{pmatrix}\,, \quad \mathbf{w} \equiv \begin{pmatrix} p_b & p_b \\ p_b & p_b \end{pmatrix}\,. \tag{123}$$

(Note that since $\sum_i \mathrm{A}_{ij} = 1$, the transformation eq. 121 conserves the norm $\|\mathbf{q}\| \equiv \sum_i q_i$.) In the preceding calculation, $k$ has been assumed to be suitably large; considering it virtually infinite, the Kirkwood correlation factor associated to the dipole oriented along $+\hat{\mathbf{x}}$ can be expressed in the language of eq. 121 as follows:

$$g_{\mathrm{K}} = \lim_{n\to\infty} g_{\mathrm{K},n}\,, \tag{124}$$

where the $n$-th element of the sequence,

$$g_{\mathrm{K},n} \equiv 1 + 2\sum_{i=1}^{n} \left[\left(\mathbf{A}^i \cdot \mathbf{q}\right)_1 - \left(\mathbf{A}^i \cdot \mathbf{q}\right)_2\right]\,, \tag{125}$$

is the correlation factor for the $(n+1)$-th dipole of an open $(2n+1)$-cluster (in the hypotheses that all dipoles in the chain are equal and that polarization mechanisms other than those accounted for through $E_b$ are negligible).

For the general $i$-th term in the series, eq. 125, we have the following identity: $\left(\mathbf{A}^i \cdot \mathbf{q}\right)_1 - \left(\mathbf{A}^i \cdot \mathbf{q}\right)_2 = (p_s - p_f)^i$, which can be

proved by mathematical induction. Then, the Kirkwood factor can be found from the summation of a geometric series:

$$g_{\mathrm{K}} = \frac{1 + p_s - p_f}{1 - p_s + p_f} \,. \tag{126}$$

Note that $g_{\mathrm{K}} = 1$ if $p_s = p_f$, i.e. like in the freely jointed chain case; moreover, eq. 126 does not carry information anymore about the orientation of the fixed dipole with respect to which $g_{\mathrm{K}}$ itself has been calculated.

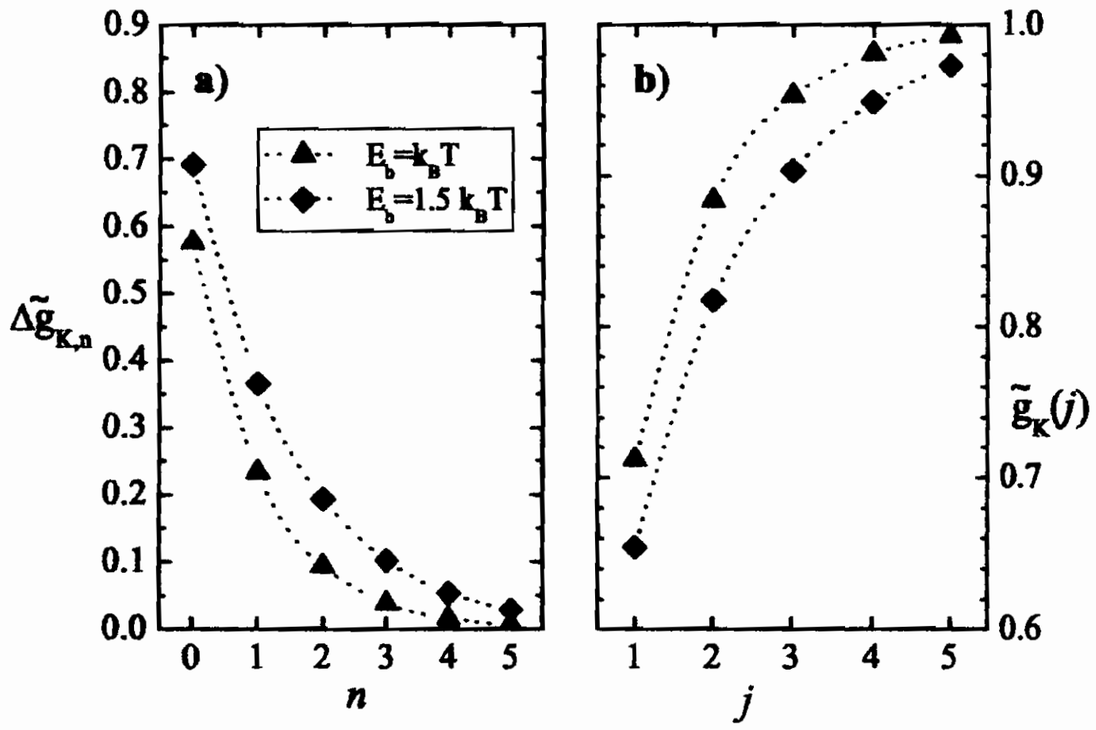



Figure 29: Convergence to the limit of $\Delta\tilde{g}_{\mathrm{K},n}$, eq. 127, and of the normalized quantity $\tilde{g}_{\mathrm{K}}(j) \equiv g_{\mathrm{K}}(j)/g_{\mathrm{K}}$, which is the correlation factor of the $j$-th dipole of a semi-infinite chain. The energy values are appropriate to realistic situations at room temperature.

Figure 29 reports the convergence to the limit of the normalized deviation

$$\Delta\tilde{g}_{\mathrm{K},n} \equiv 1 - \frac{g_{\mathrm{K},n}}{g_{\mathrm{K}}} \,. \tag{127}$$

The dependence of the normalized correlation factor of the $j$-th dipole of a semi-infinite chain, as a function of its position with respect to the end, is shown in the same figure. Note the fast convergence in both cases; this suggests that assigning the same

$g_K$ to any of the dipoles is often a good approximation. This is in fact the case for long chains, but for diluted alcohols things may be different although not trivial for a quantitative assessment.

### 3.3.3 Fitting the data

Once the expression of the correlation factor $g_K$ has been derived in terms of the phenomenological energies $E_\nu$, we can turn back to eq. 117 and use it to analyze the results of dielectric constant measurements carried out at different temperatures. It is evident that not only chain stiffness (i.e. $g_K$) plays a role, but also our first approximate model, because we need it to estimate $\eta$. One more important point has to be looked at, namely that the fitting to the data should not lead to values of $E_\nu$ close to, or larger than the hydrogen bonding energy $E_h$. Should this happen, it would mean that the model doesn't hold. The freely jointed chain approximation assumes indeed that the energy associated to a bond is roughly $E_h$, independent of the bond state. An exception is for the *fold* state, but this is already known to be practically ruled out by the condition that no multiple occupancy of any lattice site is allowed.

Besides, we must be aware of the fact that in our theory the alcohol molecules are structureless, so we may expect that some difficulty could arise when applying the model to molecules with alkyl tails which are long or of complicated shape.

In order to estimate $\eta$ as a function of the temperature, we use the value $E_h = 3.5$ kcal/mol obtained independently by means of the dielectric analysis of diluted alcohol-in-$CCl_4$ solutions.

If $v_a$, $\varepsilon$ and $\varepsilon_{\infty,a}$ are known as functions of the temperature, we can fit the data using $\mu_0$ and $g_K$ as variable parameters.

It is not the case to reproduce the whole procedure here (see ref. [21] for the details and also for the data), however the results of the analysis carried out on different alcohol species are reported in Table 2.

The complete form of $g_K$, eq. 126, led to values of $E_f$ rather

| | $\mu_0$ (D) | $E_b$ (kcal/mol) | $g_K$ ($T = 298$ K) | $\Delta T$ (K) |
|---|---|---|---|---|
| methanol | 2.19 | 0.58 | 2.34 | 230-398 |
| ethanol | 2.16 | 0.62 | 2.42 | 195-303 |
| 1-propanol | 2.12 | 0.68 | 2.58 | 184-298 |
| 1-butanol | 1.96 | 0.85 | 3.1 | 290-320 |
| 1-pentanol | 1.81 | 0.94 | 3.47 | 290-320 |
| 1-hexanol | 1.66 | 1.09 | 4.17 | 290-320 |

Table 2: Molecular dipole $\mu_0$ and bending energy $E_b$ fitting the data of the different alcohol species.

larger than $E_b$; so, all the data could be suitably analyzed with the following approximate and more compact expression:

$$g_K \simeq \frac{Z_c + 1}{Z_c - 1}. \tag{128}$$

The temperature dependencies of $\mu_0^2 g_K$ for methanol and 1-pentanol are reported in Fig. 30.

From the data reported in the table we can see a systematic increase of the bending energy $E_b$ with the length of the alkyl tail. This would seem to be reasonable, since longer tails would imply a better efficiency of hindering mechanisms (thus $E_b$ would include some entropy contribution implicitly). However, we cannot trust this interpretation at the outset, because we didn't envisage any role of the alkyl tail in our model. Also the fact that the fitting value of the molecular dipole correspondingly decreases is a sign that maybe something is being missed, also because it is difficult to interpret. By the way, the $E_b$ values are reasonably lower than $E_h$ and, apart of the heavier molecules, witness for some reliability of the model.

Most important are the fitting results for $\mu_0$, because they are a measure of the dipole moment of these alcohols *in the liquid phase.*

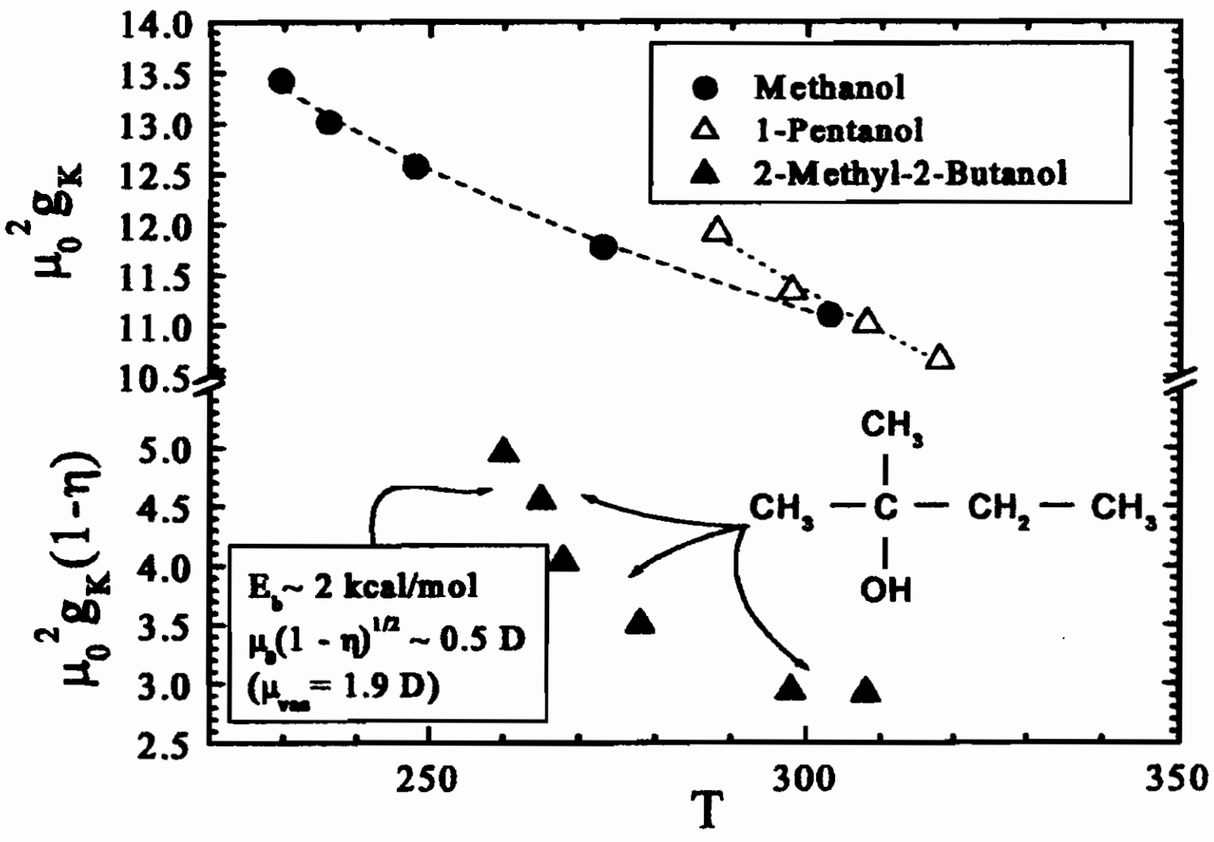



Figure 30: The factor $\mu_0^2 g_K$ (in $D^2$) for methanol and 1-pentanol as a function of the temperature. The case of 2-methyl-2-butanol is also shown for a comparison, but since the freely jointed chain model does not work, the quantity $\mu_0^2 g_K(1-\eta)$ is reported instead. Note the exceedingly high value of $E_b$.

Since now (to the knowledge of the author), the dipole moment of these molecules was indeed measured only in their gas phase, with a resulting value of $\sim 1.7$ D for all of them. Our only support to the reliability of this result is represented by numerical simulations (on methanol and ethanol), with which our numbers agree quite well [22].

Figure 30 also shows the experimental data for the quantity $\mu_0^2 g_K(1-\eta)$ for 2-methyl-2-butanol. The resulting fitting value of $E_b$ points out a serious inadequacy of the model in this case. The reason is that 2-methyl-2-butanol has a more complicated shape than its linear counterpart (1-pentanol).

### 3.3.4 Heterogeneous association

If we examine Fig. 23 with a bit more attention, we see that in the case of methanol the dielectric constant is larger than ideal in the high-$\phi$ region. This experimental behavior is interesting, because if we extrapolate to the intercept at $\phi = 0$ the ideal behavior fitting the high-$\phi$ data, we find that *things are as if* $CCl_4$ has an apparent (ideal) bulk dielectric constant of $\varepsilon_s \simeq 20$ (see Fig.31). This value can't be realistic for a non polar molecule such as $CCl_4$; on the other hand, in the high-$\phi$ region it is unlikely that there is any bulk $CCl_4$ at all. We envisage the possibility that there are two possibly

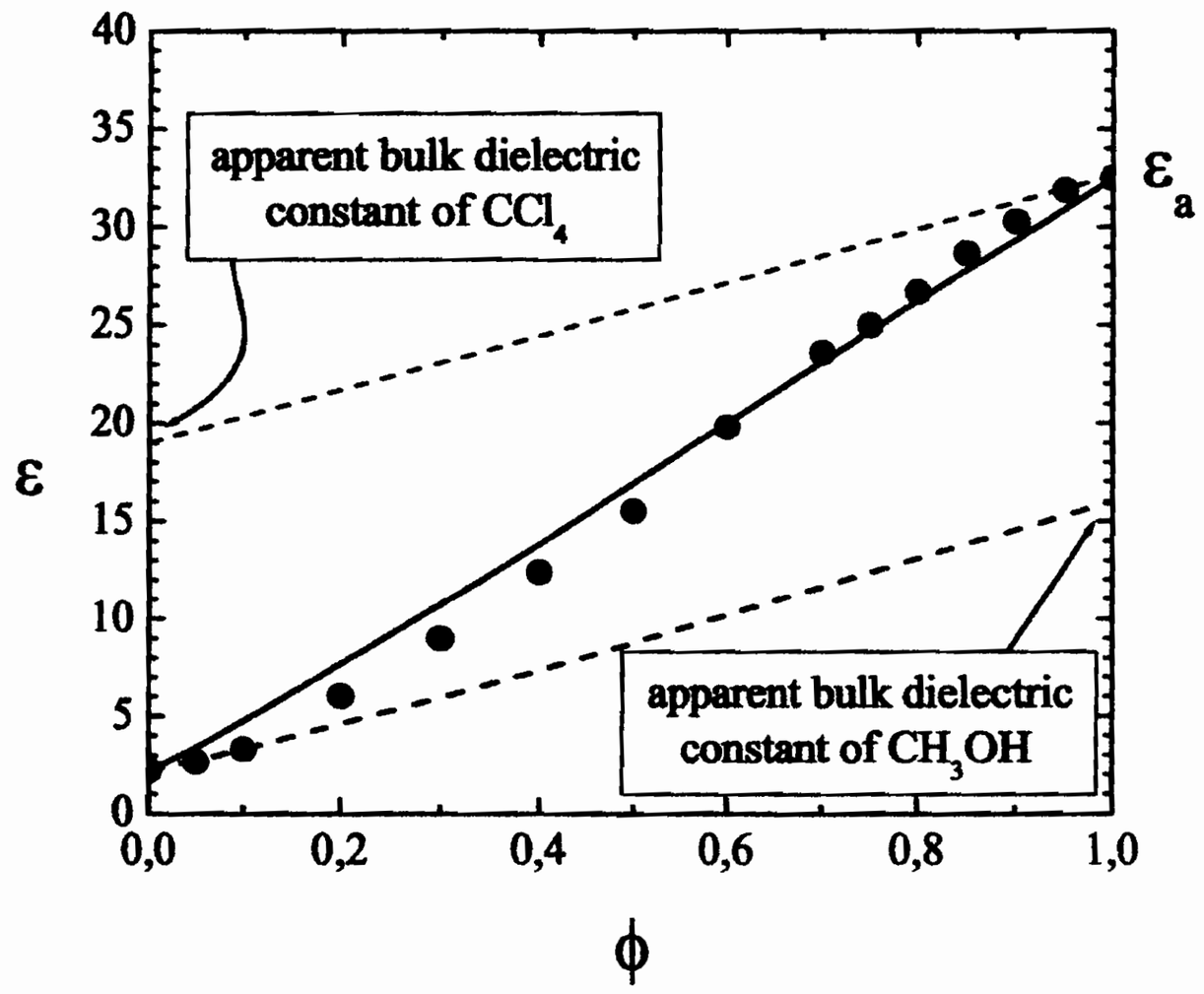



Figure 31: Dielectric constant at room temperature as a function of the alcohol volume fraction for the methanol/$CCl_4$ mixtures. The lines represent ideal behaviors.

concurring reasons for this peculiar behavior, namely:

1. $CCl_4$ associates with the chain end methanol (this has been also monitored by calorimetry through the enthalpy of mixing

at high $\phi$ [16]); this process causes a significant decrease of the average number of cyclic clusters in the alcohol component.

2. associated $CCl_4$ acquires a quasi-permanent dipole moment $\boldsymbol{\mu}_s$, the lifetime of which is the same as that of the associating bond.

We shall refer to association between $CCl_4$ and methanol as to the *heterogeneous association.* The above points are just a suggestion, thus susceptible of confutation. Keeping this in mind, we have to assess whether these hypotheses are a reasonable basis for further developments, or a nonsense at the outset.

To this aim we first have to change eq. 95 properly. Our two issues imply the following:

$$\frac{9k_BT}{4\pi\varepsilon}\left\{\frac{\varepsilon-\varepsilon_{\infty,a}}{2\varepsilon+\varepsilon_{\infty,a}}\phi+\frac{\varepsilon-\varepsilon_{\infty,s}}{2\varepsilon+\varepsilon_{\infty,s}}(1-\phi)\right\}= \frac{N_A}{v_a}\left(\frac{\varepsilon_{\infty,a}+2}{2\varepsilon+\varepsilon_{\infty,a}}\right)^2\mu_0^2 g_{\mathrm{K}}(1-\eta)\phi+ \frac{N_A}{v_s}\left(\frac{\varepsilon_{\infty,s}+2}{2\varepsilon+\varepsilon_{\infty,s}}\right)^2\langle\boldsymbol{\mu}_s\cdot\mathbf{M}\rangle_0(1-\phi) \tag{129}$$

where the first term refers to methanol and accounts for cyclic clusters through the factor $(1-\eta)$, while the reminder describes the presence of a new polar component, namely, the associated $CCl_4$ with its induced dipole $\boldsymbol{\mu}_s$; $v_s$ is the molar volume of $CCl_4$. (To make things simple, we assume that *all* $CCl_4$ associates with methanol at high enough $\phi$.)

Point (1) relates to the change of $\eta$ caused by the fact that the presence of the associating (solvating) $CCl_4$ affects the overall average hydrogen bonding energy $E_h$. Indeed, let $E_{h0}$ be the energy needed to separate two hydrogen bonded molecules in the absence of any solvent. When the bond rupture takes place in the neighborhood of an isolated, non-associated $CCl_4$ molecule, a subsequent association process with an end-methanol would be accompanied by

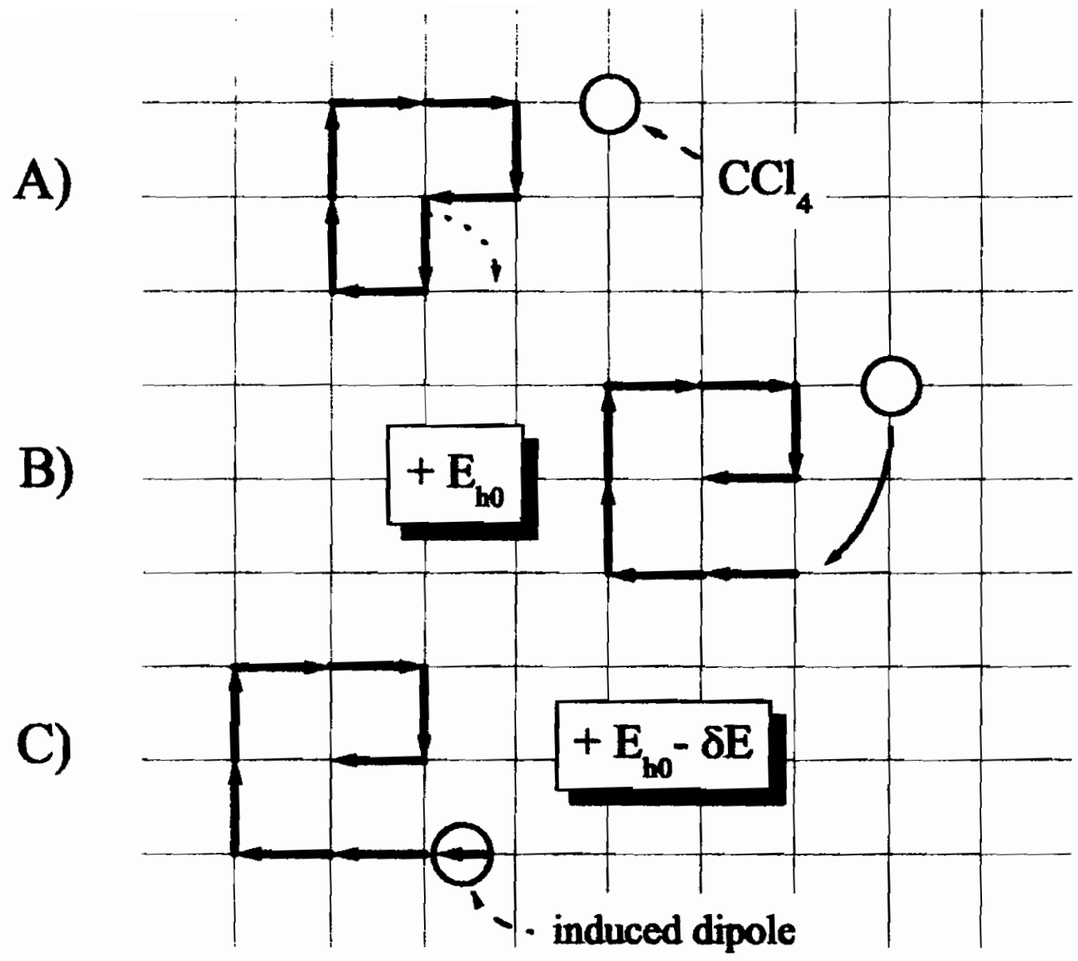



Figure 32: Rupture of a chain and heterogeneous association. The $CCl_4$ molecule acquires a quasi-permanent dipole moment due to association.

an energy gain $\delta E$, and the final overall energy involved in the bond rupture would be $E_{h0} - \delta E$ (see Fig. 32).

If the number of $CCl_4$ molecules increases, the average bonding energy $E_h$ decreases from $E_{h0}$ down to a saturation value $E_{h0} - 2\delta E$, when both the ends of each chain are associated with the solvent. The number of chain-end moles is

$$N_{end} \simeq \frac{2N(1-\eta)}{\langle n_o \rangle}\,, \tag{130}$$

where $N = \phi V/v_a$ is the total number of alcohol moles. On the other hand, the number of moles of $CCl_4$ is $N_s = (1-\phi)V/v_s$, so that at the end we find, before saturation is reached,

$$E_h \simeq E_{h0} - 2\delta E\,\frac{N_s}{N_{end}} = E_{h0} - \delta E\,\frac{(1-\phi)v_a}{\phi v_s}\,\frac{\langle n_o \rangle}{1-\eta}\,. \tag{131}$$

For a consistent calculation, the values of $\langle n_o \rangle$ and $\eta$ entering eq. 131 must coincide with those resulting from the equilibrium distribution obtained by taking the energy $E_h$ itself for the calculation of the distribution. An iteration procedure allows for the calculation of these quantities as functions of the "parameters" $E_{h0}$, $\delta E$ and $\phi$.

About point (2), we have the great advantage that the magnitudes of the methanol dipole and of the correlation factors are known (see Table 2). We estimate the average $\langle \boldsymbol{\mu}_s \cdot \mathbf{M} \rangle_0$ by assuming that the induced dipole of a $CCl_4$ molecule is oriented along the same direction of the chain-end methanol dipole to which it is associated. Only a semi-infinite methanol chain contributes to the correlation factor, and considering that $g_K$ refers to an "inner" dipole (see eqs. 124 and 125), the following approximation is adopted:

$$\langle \boldsymbol{\mu}_s \cdot \mathbf{M} \rangle_0 \approx \mu_s \left( \mu_s + \mu_0 \frac{g_K + 1}{2} \right) . \tag{132}$$

With reference to a practical situation, the actual value of the dielectric constant for the methanol/$CCl_4$ mixture is $\varepsilon = 31.9$ at $\phi = 0.95$. On the other hand, assuming for example $E_{h0} = 3.5$ kcal/mol and $\delta E \sim k_B T$ at room temperature [23], a fixed point in the calculation of $\langle n_o \rangle$ and $\eta$ is found with $E_h \approx 3.3$ kcal/mol (in working out this number, it has also been taken into account that only the appropriate Boltzmann fraction of solvent molecules undergoes heterogeneous association even if $N_{end} > N_s$ : a detail not mentioned above for the sake of simplicity).

By imposing that $\varepsilon$ as calculated from eq. 129 must be equal to the actual experimental value, we find that the dipole induced by heterogeneous association on $CCl_4$ must have an approximate magnitude of $\mu_s \simeq 1.7$ D.

### 3.3.5 Outlook

So, at the end we developed a mean field theory for the analysis of complicated systems. We contracted all long range interactions to merge together with the short range ones characterizing a single

bond. This way we have hidden a lot of problems behind a *phenomenological* fitting parameter, i.e. $E_b$. The latter carries with it some information on the long range interactions, because when we fit the data we take also them altogether.

With regards to the Markov assumption, this is an important issue which deserves further investigations. because there is some intrinsic contradiction into our model which could be solved. To be more explicit, this hypothesis enters also our freely jointed chain model (for us the chains are ideal), and conflicts with the constraint that $V$ and $\phi$ are held fixed when the equilibrium cluster distribution is to be found; in other words, on the one hand we allow multiple occupancy and calculate the configurational entropy consequently, on the other we put a constraint relating to the impossibility of multiple occupancy while keeping the same expression for $W_t$.

Moreover, the Markov assumption again, together with the fact we take structureless molecules, does not allow for the role of steric hindrance to be pointed out by our analysis, and consequently to disentangle the entropy contribution from $E_b$. We can understand that this last point has to do with the fact that ideality in the low-$\phi$ region persists up to $\phi \simeq 0.3$ in the case of 1-pentanol, while just up to $\phi \simeq 0.1$ in the case of methanol (see Figs. 23 and 31). Probably, in the case of 1-pentanol we really are in the presence of an almost ideal mixture made of $CCl_4$ on the one hand, and of small alcohol clusters on the other. Thus, in the case where the tail is longer, there seems to be a tendency to form many small clusters instead of few larger ones. Since the preference of a group of molecules to associate in large or small clusters depends on the configurational entropy, it is clear that the problem of describing with more detail the structure of the alcohol molecules deserves being considered in further investigations.

# References


[1] H. B. Callen, *Thermodynamics and an introduction to thermostatistics*, 2nd ed. Wiley

L. D. Landau and E. M. Lifšits, *Fisica Statistica 1a parte*, Editori Riuniti

W. Greiner, L. Neise and H. Stöcker, *Thermodynamics and Statistical Mechanics*, Springer

R. Bowley and M. Sanchez, *Introductory Statistical Mechanics*, Oxford

D. Chandler, *Introduction to Modern Statistical Mechanics*, Oxford

[2] G. Strobl, *The Physics of Polymers* Springer

[3] M. Doi, *Introduction to Polymer Physics*, Oxford

[4] J. H. Gibbs and E. A. DiMarzio, J. Chem. Phys. **28**(3), 373 (1958).

[5] G. Adam and J. H. Gibbs, J. Chem. Phys. **43**(1), 139 (1965).

[6] P. M. Chaikin and T. C. Lubensky, *Principles of Condensed Matter Physics* Cambridge

[7] M. Pieruccini, G. Di Marco and M. Lanza, J. Appl. Phys. **80**(3), 1851 (1996)

[8] F. Aliotta, G. Di Marco and M. Pieruccini, Physica A **298**, 266 (2001)

[9] F. Aliotta, G. Di Marco, R. Ober and M. Pieruccini, J. Appl. Phys. **93**(9), 5839 (2003)

[10] G. Di Marco and M. Pieruccini, *Polymer Crystallization; Observations, Concepts and Interpretations*, J. U. Sommer and G. Reiter (Eds.), Springer Lecture Notes in Physics 2003

[11] M. Imai, K. Kaji and T. Kanaya, Macromolecules **27**, 7103 (1994)

[12] P. D. Olmsted, W. C. K. Poon, T. C. B. McLeish, N. J. Terrill and A. J. Ryan, Phys. Rev. Lett. **81**, 373 (1998)

[13] Santa Cruz et al. J. Polym. Sci. (B), 29, 819 (1991)

[14] T. Asano, F. J. Baltá Calleja, A. Flores, M. Tanigaki, M. F. Mina, C. Sawatari, H. Itagaki, H. Takahashi and I. Hatta, Polymer **40**, 6475 (1999)

[15] M. Doi and S. F. Edwards, *The Theory of Polymer Dynamics* Oxford

[16] J.-E. A. Otterstedt and R. W. Missen, Trans. Faraday Soc. **57**, 879 (1962)

[17] M. Pieruccini, F. Saija, R. Ponterio and C. Vasi, J. Chem. Phys. **119**, 10771 (2003)

[18] H. Fröhlich, *Theory of Dielectrics - Dielectric Constant and Dielectric Loss* 2nd ed. Oxford 1987

[19] R. Veldhuizen and S. W. de Leeuw, J. Chem. Phys. **105**, 2828 (1996).

[20] J. G. Kirkwood, J. Chem. Phys. **7**, 911 (1939).

[21] M. Pieruccini and F. Saija, J. Chem. Phys. **121**, 3191 (2004).

[22] M. Haughney, M. Ferrario, and I. R. McDonald, J. Phys. Chem. **91**, 4934 (1987); T. Fonseca and B. M. Ladanyi, J. Chem. Phys. **93**, 8148 (1990); J. Gao, D. Habibollazadeh, and L. Shao, J. Phys. Chem. **99**, 16460 (1995); M. E. Martin, M. L. Sanchez, F. J. Olivares del Valle, and M. A. Aguilar, J. Chem. Phys. **116**, 1613 (2002); M. Pagliai, G. Cardini, R. Righini, and V. Schettino, J. Chem. Phys. **119**, 6655 (2003).

[23] H. Torii, Chem. Phys. Lett. **393**, 153 (2004).

# Positron Emission Tomography: status of the art and future perspectives

Monica M. Necchi*
Department of Nuclear and Theoretical Physics
- University of Pavia -
via A. Bassi 6 - 27100 - PAVIA (PV)

## 1 P.E.T.: present features

Positron Emission Tomography (PET) is among the most sensitive methods to trace amounts of molecules in vivo. PET provides a kind of metabolic information that other imaging methods are unable to provide; therefore this technique is used to measure in man or in living animals biochemical and physiological processes in any organ with 3-D resolution. The last 25 years have seen a rapid and still ongoing development in the production of positron emitters, radiochemical labelling techniques, tomograph technology and image reconstruction algorithms.

## 2 The method

The use of positron emitters for radioactive labelling offers important advantages compared to single photon emitters[1]. Some of the physiologically most interesting chemical elements like Carbon, Nitrogen and Oxygen have only positron emitting short-lived isotopes. This makes a detector sensitive to positron decay highly desirable.

*Research contract PRIN no. 2003-02-9728 - MIUR

Positrons travel only a short distance in tissue[1], then annihilate with an electron into two 511 keV $\gamma$ emitted back-to-back. When these two photons are detected by two opposite detectors the decay event can be located along the line connecting the two detectors. The two 511 keV gamma rays are detected by a coincidence circuit with a narrow time window, usually about 15 ns. These detectors capture the photons emitted by the isotope from within the system under study at a total rate of up to hundreds of millions per second, while the transducers (such as PMTs) convert them to electrical signals sending them to proper electronics. A proper mathematical algorithm corrects the raw data for scatter, attenuation, accidental coincidence; it normalizes for differences in detector efficiencies and reconstructs the spatial distribution of the radioactivity density inside the organ or the system under study.

Since both photons have the same energy the detection probability is almost independent from the point between the two detectors where the annihilation occured. Therefore the corrections for attenuation in tissue, which are generally huge in these measurements, can be done very accurately.

Commercial tomographs consist of thousands of detectors, which are arranged in several parallel rings surrounding the object to be investigated. Pairs of opposite detectors are connected in coincidence. Originally this was restricted to opposing detectors in the same ring or in neighbouring rings with septa shields in between them to get rid of unwanted scattered radiation. More recently, to increase sensitivity and acceptance the septa were removed and coincidences between many rings were registered. Now the most advanced systems have no shielding at all between the detectors. The accepted solid angle is increased by extending the axial Field Of View (FOV) with more rings and tightly packed detectors with a high stopping power arranged in a compact geometry. This 3D or volume data aquisition, where all possible coincidence lines inside the detector volume are registered, induces, besides higher counting rates of true coincidences, a large increase of scatter and random coincidences. A scatter coincidence is a detected event where one or both photons were deflected by Compton scattering and are therefore assigned to wrong coincidence

[1]The $\beta^+$ emitted by radioactive fluorine annihilates with a nearby electron after a path of about 1.4 mm to over 13 mm depending on the radioisotope emploied.

lines. Scattered coincidences may reach the same amount as the true events and very elaborate correction algorithms have become necessary to subtract scatter contributions for accurate quantitative results. A random coincidence is an event where two photons from different decays are detected during the short coincidence time window. Random coincidences can be subtracted online using a second delayed coincidence window.

# 3 New developments

The improvements of the PET tomographs spatial resolution was a primary goal since the beginning [1]. Early tomographs used Tallium doped NaI (NaI(Tl)) crystals as scintillation detectors[2]. NaI(Tl) crystals were soon substituted by BGO (Bismuth Germanate) crystals which have higher detection efficiency and are not hygroscopic.
Commercial tomographs using BGO reached a spatial resolution of approximately 4 mm. PET detectors require that the detector material be of high enough Z to maximize the photoelectric cross sections for 511 keV, be fast enough to handle the high counting rates involved, and be able to separate genuine coincident events. Furthermore, detectors for PET should show a reasonable pulse height resolution for 511 keV photons so that scattered events in the material under study can be rejected.
Presently[3] the first tomographs with a new scintillator material LSO (Lutetium Oxyorthosilicate) are being built. LSO has similar high detection efficiency as BGO, but with a five fold higher light output and an eight times faster light decay time. This allows to push spatial resolution close to its physical limit of 1-2 mm, which is determined by the positron range in tissue and the small noncollinearity of the annihilation photons. Increasing spatial resolution requires smaller scintillation crystals. When the size of the PMT became the resolution limiting element, the one-to-one coupling between crystal and PMT

[2]Brownell and Sweet at Massachusetts General Hospital made the first positron medical image in 1951. The imaging device emploied two simple NaI(Tl) crystals, moved manually to scan brain tumors. In the 1960's and 1970's positron imaging devices used detector's arrays.

[3]New improvements in PET tomographs deal with crystals with better energy resolution (LSO) and shorter decay time (Gadolinium Oxyorthosilicate).

was given up and an Anger camera principle was applied in the block detector concept. For example, 8x8 crystals were cut into a single scintillator block forming a light guide to 4 PMTs. By comparing the signals of the 4 PMTs the individual crystals are uniquely identified. An even more cost effective scheme, a quadrant sharing arrangement of PMTs allows to further increase the number of crystals per PMT. This was recently achieved in the first PET tomograph for neurological studies in humans with LSO scintillators, the High Resolution Research Tomograph (HRRT), manufactured by SIEMENS/CTI, Knoxville, TN, USA and installed at the Max-Planck-Institute for Neurological Research in Koln. It consists of 120,000 single crystal elements, 2.1 mm x 2.1 mm each, arranged in 104 rings giving a reconstructed spatial resolution of less than 2.5 mm in a 20 cm diameter volume. This made it also necessary to measure the depth of interaction (DOI) of the incident photons in the 15 mm deep detectors to avoid ambiguities when photons penetrate several crystals by oblique incidence. From several ideas how to obtain DOI information, a scintillation phoswich with two crystal layers with different light decay times was chosen. Pulse shape discrimination allows then to separate scintillation events in the two layers.

After correction for scatter and random events, attenuation and dead time an image of the activity distribution can be reconstructed. Originally this was done with standard 2D filtered backprojection algorithms giving a stack of transverse image slices. With 3D data aquisition this was extended to volume reconstruction, which required hours of computing time. Therefore the algorithms were parallelized and run on multiple processors. Only recently, with the rapid developments in computer hardware, iterative 3D reconstruction is becoming routinely feasible.

Most recent trend in industrialized countries is the use of hybrid PET/CT systems[9], which combines a dedicated PET with an X-ray computerized tomography (CT) scanner in the same instrument. The CT images provide a map for PET attenuation correction and an anatomic framework for the PET metabolic information.

# 4 Future trends for PET tomographs

PET detector development progressed significantly since the first PET tomographs. The number of elements increased from a few ten elements in the first tomograph at St. Louis Hospital to more than 100,000 in the new HRRT that has been delivered to Max-Planck-Institute in Koln on February 1999[2]. The HRRT is the first commercial LSO tomograph and it represents a major change in the detector technology for PET. The ECAT HR+ at that time was the highest resolution commercial PET tomograph, using BGO detector crystals. Commercial tomograph with the same detector as the ECAT HR+ but having six rows of detectors rather than four is the ECAT HR++, delivered to the MRC Hammersmith Hospital in London in 1997. The ECAT HR++ has 27,648 detector elements. These BGO tomographs have the highest feasible density detector elements and use large numbers of expensive photomultipliers (1728 for the HR++). Since BGO has limited light output and relatively slow decay the pratical limit in image resolution is reached for these tomographs.
LSO allows the use of fewer PMTs for the PET detectors and at the same time increases the perfomance in counting rate, sensitivity and image resolution. As is evident from data in Table 1, LSO has allowed

| | PET/ SPECT | ECAT HR+ | ECAT HRRT |
|---|---|---|---|
| Sensitivity ($\cdot 10^3$) | 850 | 1200 | 4000 |
| Image resolution | 4.5 mm | 4.5 mm | 2.5 mm |
| Peak true count rate | 100,000 | 200,000 | 900,000 |
| number of PMTs | 198 | 1152 | 1120 |

Table 1: Performance comparisons of three tomographs

an improvement of the HRRT over the HR+ of more than a factor three in sensitivity, almost a factor of two in space resolution, a factor of 4.5 in count rate and still the HRRT has fewer PMTs. This is a breakthrough in detector technology not only for research tomographs, but also for clinical tomographs.

## 4.1 Scintillator choice

NaI detectors operated in Anger logic mode have served the gamma camera market well in the past, although the hygroscopic nature of the detecor material does not favour the development of such desirable features as crystal segmentation. This feature is indeed valuable when the detector is operated without a collimator and thus requires high count rate and good linearity performances. The nuclear market trend toward Fluorine Deoxy Glucose (FDG) imaging creates the need to acquire images at 511 keV in a reasonable time, which is very inefficent with current NaI detectors. On the other hand, the current BGO detector for PET has insufficient light output to provide adequate energy resolution for the 140 keV currently used in SPECT. Both BGO and NaI have relatively long light decay time (300 ns and 230 ns, respectively), which may limit their counting rate capability. The technology breakthrough in positron imaging is the discovery of LSO crystal which has a number of advantages over the other scintillators used for nuclear medicine instrumentation[4]. Table 2 compares the main parameters of NaI, BGO, $BaF_2$, LSO crystals. For 511 keV

| Parameter | NaI | BGO | $BaF_2$ | LSO |
|---|---|---|---|---|
| Density ($g/cm^3$) | 3.7 | 7.1 | 4.9 | 7.4 |
| Mean free path (cm at 511 keV) | 2.9 | 1.1 | 2.2 | 1.2 |
| Index of refraction | 1.85 | 2.15 | 1.50-1.54 | 1.82 |
| Hygroscopic? | yes | no | no | no |
| Decay time (ns) | 230 | 300 | 0.8-630 | 40 |
| Light output [NaI(Tl) = 100] | 100 | 15 | 5-21 | 75 |

Table 2: Properties of NaI(Tl), BGO, $BaF_2$ and LSO for 511 keV photons

photons, LSO has a mean free path almost equal to that of BGO and is thus an as efficient scintillator with the same potential for high resolution imaging. The light output of LSO is three-fourths of NaI, thus much larger than that of BGO or $BaF_2$, resulting in a much better energy resolution. The LSO crystal is very rugged and is non-hygroscopic. The scintillation light decay time for LSO is 40 ns compared to 230 ns for NaI, 300 ns for BGO and 0.8 (630) ns for the fast (slow) scintillation process of $BaF_2$. It provides a

rapid time response and may give a time resolution of 750 ps FWHM. Thus a coincidence time window of about 4 to 6 ns is feasible, which should significantly reduce random coincidences relative to current PET cameras, and improves the signal-to-noise ratio. Moreover the event processing time for LSO can be almost six times shorter than that achievable with BGO or NaI, allowing better counting rate capabilities.

# 5 Tracer production

The commonly used positron emitting nuclides in PET are $^{11}C$,$^{13}N$,$^{15}O$ and $^{18}F$. The first three isotopes are the most abundant elements in organic compounds[1]. This allows the labelling of naturally occuring biomolecules or drugs without altering the compounds defined courses in living systems. The $^{18}F$ isotope can be used to substitute hydrogen or hydroxyl groups. Because of their short half-life, ranging from 2 min for $^{15}O$ to 109 min for $^{18}F$, they have to be produced close to their application. Various small, compact cyclotrons specially built for the production of short lived isotopes in a hospital environment are commercially available. There are cyclotrons with external proton and deuteron beams, self-shielding machines with internal targets down to 3 MeV beam energy for the exclusive generation of $^{15}O$ and negative ion machines allowing simultaneous irradiations on separate target ports. Computer controlled automated robot systems synthetize in heavely shielded hot cells radiolabelled compounds in optimized short times of few minutes, including purity controls.

The most commonly used PET radiotracer is FDG, a glucose analog, metabolic imaging agent giving precise and regional information of energy metabolism in brain, heart, other organs and tumors[9]: its uptake in tissue can be easily quantified as glucose metabolic rate. Biochemical processes of the body's tissue, such as metabolism of glucose, are altered in virtually all diseases and PET detects these changes by identifying areas of abnormal metabolism, which are indicated by high photon emission. Cancer cells, for instance, typically have a much higher metabolic rate, because they are growing faster than normal cells, thus they absorb 60 to 70 times more sugar than

normal cells and consequently emit more photons.
The production of $^{18}F$ used for PET is based on the nuclear reaction $^{18}O(p,n)^{18}F$, where water with 98% of $H_2^{18}O$ as target material is employed; the reaction produces 800 mCi, corresponding to 32 GBq[3]. The relatively high half-life (109 minutes)[9] of $^{18}F$ allows the radiotracer distribution to local hospitals after its production in a regional center. The half-life of 109 minutes also provides an appropriate decay time for searching tumors and metastases by patient whole body scans.
The high uptake of FDG by the tumor produces very high-localized intensity PET images in which is very difficult to see other organs and structures. To evaluate the exact location of a tumor for staging a patient, for treatment or follow-up planning, it is necessary to setup an anatomic framework for the metabolic information provided by PET. In addition, for better and more exact quantification of FDG uptake, an anatomical map of the structures imaged by PET is needed. These two reasons lead to the development of hybrid PET/CT systems in which both are combined in one instrument. By combining PET and CT in the same instrument, images of both approaches are automatically registered. PET images have an anatomical reference and photon attenuation can be corrected with higher accuracy[9].

# 6 Complete Body Screening (3D-CBS)

Dario Crosetto is the inventor of the 3D-CBS system and founder of 3D-Computing[6]. Crosetto studied how to improve, by over 400 times, the efficiency of the current PET machine for whole-body examination. These efficiency improvements are made possible because of the novel electronics and detector assembly of the 3D-CBS machine, which is integrated with a CT scan. The 3D-CBS allows the use of a larger area of economical crystal detectors thereby exposing the patients to approximately 4% of the radiation they currently receive.
Before the advent of this invention, screening of the entire body was not advisable because current PET machines expose the patients to over 10 times the radiation recommended by the International Commission of Radiation Protection. This high dose is required,

because most of current PET tomographs detect with less accurate measurements only one out of 10,000 photons emitted from the patients body. In addition, current PET examinations are slow and cost prohibitive.

Crosetto's improvements stem from an innovative way to detect more accurately a larger number of photons emitted by the tracer medium than current designs can do. The patent-pending device will allow for detection of more photons, more accurately, thus reducing radiation doses to the patient, and improving the image quality. Furthermore, it allows the examination of more patients per hour, therefore reducing costs.

The key features of his invention are supported by simulations and hardware implementations. Crosetto's radical improvements in PET efficiency are supported by the new architecture of his electronics using a set of DSP (Digital Signal Processor) on each electronic channel providing the capability to exchange information received from neighboring detector elements and to execute complex algorithms that can measure more accurately the total energy and the spatial resolution of the incident photon. It eliminates also the parallax error of the oblique photons, allowing reduction of false positives, false negatives and an increase in image sharpness. Another supporting feature is the way the signals from the detectors are connected to the set of DSPs on each electronic channel and the new way the detector is assembled as a single (or few) camera(s) made of hundreds of sensors, each capable of finding a photon candidate, versus current PET, which is an assembly of hundreds of small cameras, each with lower energy and spatial resolution at the edges and corners with respect to the center.

The innovations described above allow increasing the length of the PET detector, using economical crystals or different kinds of cheaper detectors, from the current 16 cm to over one meter (when the actual length of the detector is doubled, the number of photons captured is increased by a factor of four).

Untill a few years ago there was the belief that the only way to improve PET efficiency was to improve the efficiency of the crystal detector and not to improve the electronics. During the past 25 years, improvements in PET design have achieved higher efficiency by a factor of only 2 to 3 times every 5 years.

Crosetto's novel technique shows that PET efficiency can be improved by using a special massively parallel-processing system with digital signal processing on each electronic channel. None of the new PET components (crystals, PMT, Avalanche Photo Diode or the new Flat Panel Sensors that replace the PMT) can improve PET efficiency by more than a factor two or so.
The 3D-CBS higher sensitivity will more effectively show abnormal biological processes at the molecular level, before the cancer exhibits sympthoms and before an anatomical change occurs in the body tissue, which is normally detected by CT. Past experience suggests that earlier detection achieved with regular screening using very sensitive devices can dramatically improve survival rates.

# 7 Current PET limitations

The electronics of the current PET limits its performance; it is not fully capable of extracting the complete characteristics of the interaction between the photon and the detector from signals arriving at high data input rates from thousands of sensors[5]. The electronics has been the main impediment to extending the axial FOV; the increases in efficiency that would justify extending the axial FOV are not possible with the electronics of the current PET. The current PET electronics inefficiencies in detecting photons occur because there is no independent digital signal processing at each electronic channel and there is no communication between adjacent electronic channels. This limitation affects sensitivity and spatial resolution. Sensitivity is lowered when photons striking a crystal coupled to the border of two sensors, causing them to release half (or less) of their energy in two (or more) adjacent electronic channels, are not recognized as photons because each channel receives less than the nominal energy to be considered as a 511 keV photon of a PET emission event.
Spatial resolution suffers at the edge of each 2x2 PMT block because the centroid algorithm cannot weight the PMT signals from both sides of the PMT closest to the point at which the photon struck the crystal. This causes a reduction of the overall sensivity, which translates into greater patient exposure to radiation, poorer image quality and longer scanning time.

The image quality of the current PET is poor because it has:

1. a short FOV, limited by an inefficient electronics that does not offset the cost of the detector if the FOV were increased;

2. no accurate time-stamp assigned to each photon limiting the detection of neighboring photons emitted within a short time interval, causing long dead-time of the electronics and increasing randoms (i. e. photons in time coincidence belonging to two different events), most PETs do not have any photon time-stamp assignment;

3. analog signal processing on the front-end electronics limiting photon identification because of poor extraction of the characteristics of the incident photon and no capability to improve the signal-to-noise (S/N) ratio;

4. detector boundary limitations to 2x2 PMT blocks, no correlation between signals from neighboring detector blocks, no full energy reconstruction of the photons that hit the detector; most of the current PET do not attempt to make any energy reconstruction of the event, but make decisions in accepting or rejecting a photon first and later an event based on the threshold of a single signal.

5. dead-time of the electronics. Dead-time of the electronics is due to any bottleneck (e.g. data multiplexing from many lines to a single line, input saturation, processing, output saturation) present at any stage of the electronics;

6. saturation of the electronics at the input stage due to its inability to detect and process two nearby photons that hit the detector within a short time interval.

7. saturation of the electronics at the output stage due to the limiting architecture of the coincidence detection circuit.

8. a high number of randoms due to the non accurate measurement of the photon arrival time and to the long (about 12 ns) time window used when determining if two photons belong to the same event.

9. poor measurement of the attenuation by different tissues at different locations in a patient's body. These measurements are necessary to calculate the attenuation correction coefficients for PET scans.

Furthermore, high radiation dose delivered to the patient is required by the current PET because each examination needs to capture the amount of photons which provide a sufficient statistics to yield a good image. The short FOV and the inefficient electronics allow accumulation of fewer than 2 photons in coincidence for every 10,000 emitted. This inefficiency requires one to administer a necessarily high radiation dose to the patient in order to keep the examination time within an hour.

# 8 How to improve current PET efficiency

Before attempting to improve any system, it is necessary to determine where the inefficiencies are, how large they are, how they can be reduced and by how much.

Crosetto considered a total body PET scan where the radiotracer used is water with $^{15}O$[8].

The initial number of photon pairs emitted per second by the tracer in the patient body (1424 million) and the number of photon pairs per second captured by current PET (0.2 million) are not in question, because those quantities have been measured by hospitals and universities and are in agreement with measurements done by PET manufacturers.

Of the initial quantity of photon pairs emitted from within the body, some 1210 million pairs per second are scattered or are absorbed by the body. This quantity of capturable photons leads to 15% efficiency for the first stage.

Due to the short FOV (length of the detector), another 196 million photon pairs per second, which emanate from the part of the body not covered by the detector, are lost in the second stage. This leaves only 18 million photon pairs per second remaining to be captured. This quantity is estimated by dividing the length of the FOV of the detector (16 cm) by the average length of the human body (180 cm), yielding an efficiency figure equivalent to 8.5% for this stage applying

this percentage to the 214 million capturable pairs per second leaving the body.
Some photons from within the detector area are also lost. Some others emanating from the part of the body covered by the detector leave the body at angles that escape the detector through the openings between the detector segments. This number can be calculated as a percentage of the perimeter of a circle drawn around the lengthwise cross section of the entire detector not covered by the two 16 cm segments. Thus, (178-32)/178 = 0.82, i. e. 82% of the 18 million photons pair per second remain to be captured after the third stage, which is equivalent to a total 18% efficiency for this stage.
Detector crystals do not have perfect stopping power and do not capture every photon in range. Crystals have efficiency ranging between 80% and 95%. Therefore, applying such rate to the remaining 3.2 million pairs per second still capturable after the third stage in some low cost crystal detectors, 20% are lost, or 0.65 million, and 80% remain potentially capturable, i. e. 2.5 million pairs per second. Current PET capture only 0.2 million pairs per second of the original 1424 million photon pairs per second emitted by tracer within the patient's body. Of the 2.5 million photon pair per second remaining after the fourth stage, the loss of 2.4 million pairs per second is accounted for by deficiencies in the electronics and the detector design. The efficiency of stage 5 and 6 can be calculated as equivalent to 8%, as derived by subtraction from the total inefficiency and the sum of all other inefficiencies.
It is obvious from this analysis that the section needing serious study and improvements is the last one, which provides only 8% efficiency. the first stage has the efficiency related to a natural phenomenon that cannot be changed. The second and third stage can be increased in length and solid angle only if the electronics of stage 5 and 6 are not overwhelmed by the increasing of the data acquisition rate. Stage 4, although the one in which much effort and money has been invested during the past decades, can only be improved from about 80% to something over 95% (among the so called ideal crystals, that is LSO).

# 9 The 3D-Flow

Crosetto's key innovation on the electronics (on which the 3D-Flow$^{TM}$ is based) consists in a parallel-processing architecture enabling execution of a complex real-time algorithm, calculating different types of depth of interaction with zero dead time and with data exchange with neighboring processors for a time interval longer than the time interval between two consecutive input signals[7].
This provides better energy measurement, helps to reject scatter events more efficiently and provides a way to improve spatial resolution by measuring more accurately the location where the incident photon hits the detector. It also increases sensitivity in accepting oblique photons by eliminating the parallax error accurately measuring the interaction depth.
Another aspect of the invention is the way these concepts are implemented in hardware.
The 3D-Flow$^{TM}$ parallel-processing architecture allows the execution of complex algorithms with neighboring signals correlation in real time and provides the capability to extract more accurate information from the signal generated by the interaction between the incident photon and any type of crystal detector[8]. This allows a more efficient use of economical crystals. The coupling of the detector to the electronics is made in such a way that there are no boundaries or fixed detector segmentations; rather, each sensor of the detector (PMT, Avalanche Photo Diodes, etc.) is an element of a large array with the capability to act as the center of a cluster of elements, all providing information. Finally, the one-to-one mapping of the detector array with a single array of electronic processing channels remedies the inefficiency of current PET in capturing fewer photons, less accurately at the edge and corner of each of the hundreds (or thousands) of small cameras or at the edge of detectors with fixed segmentation.
A data aquisition and processing board has been developed for high efficiency photon detection in PET/CT. The board includes 68 3D-Flow$^{TM}$ processors, each capable of executing up to 26 operations in a single cycle. The 3D-Flow$^{TM}$ DAQ IBM PC photon-detection board has the capability to execute different real time algorithms for photon detection and can be interfaced to different types of daughter analog-to-digital boards. The daughter boards provide signals carrying the

amplitude information from an Analog to Digital Converter and the time information from an analog or digital Constant Fraction Discriminator (CFD) for 16 detector channels, coupled to different types of crystals (e.g. slow: NaI(Tl), BGO, or fast: LSO, Gadolinium Oxyorthosilicate (GSO), etc.).

## 9.1 3D-Flow$^{TM}$ parallel-processing architecture

This architecture takes the parallelization process one step further than DSP and its software tools allow creation, in only a few hours, of a new application with different algorithms executed on thousands of processors. Each of the 3D-Flow$^{TM}$ processors of one layer of the 3D-Flow$^{TM}$ stack[4] executes in parallel the real time algorithm, from beginning to end, on data received from the PET detector, while processors at different layers of the 3D-Flow$^{TM}$ stack operate from beginning to end on different sets of data received from the PET detector.

The extension by the 3D-Flow$^{TM}$ architecture of the execution time in a pipeline stage beyond the time interval between two consecutive input data is illustrated by the following example: an identical circuit (a 3D-Flow$^{TM}$ processor) is copied four times: A, B, C, D; the number of times the circuit is copied corresponds to the ratio between the algorithm execution time and the time interval between two consecutive input data. A bypass switch coupled to each processor in each 3D-Flow$^{TM}$ in layer A sends one data raw to its processor and passes along to the next layer three input data packets and one output result from its processor. The bypass switches on the 3D-Flow$^{TM}$ processors at layer B send two input data packets along to the next layer, one output result received from layer A and one result from its processor, and so on. Only the processors at layer A are connected to the PET detector and these receive only input data. The processors at layer D send out only results. This architecture simplifies the connection in a parallel processing system and does not require a high fan-out from the detector electronics to send data to different processors of a parallel-processing system. All connections

[4]A stack is a set of several layers, assembled one adjacent to another to make a system.

are point-to-point with several advantages in low power consumption and signal integrity.

# 10 Conclusions

In conclusion, major developments in PET technology have already played a major part in defining and establishing the role of this imaging modality in oncology. The introduction of PET/CT is the process of refining this role. New software developments combined with the introduction of new scintillators and PET detector designs hold the potential to improve the throughput of this technique and open the way to new clinical applications.
In the present project we plan to tackle the issue of the cost of the electron detectors substituting the expensive small crystals with Multigap Resistive Plate Chambers (MRPCs) which are cheaper gas detectors and can be built and assembled in large dimensions.

# Acknowledgements

I would like to thank all the colleagues that have contributed in several ways to this work and the *Ministero dell'Istruzione, dell'Università e della Ricerca* for the financial support.

# References

[1] K. Wienhard *Positron Emission Tomography.*
[2] R. Nutt and C. L. Melcher *Current and future developments with LSO, a scintillator with excellent characteristics for PET*, Revue de l'ACOMEN, 1999, vol. 5, n$^o$ 2.
[3] *PET in more detail: Radionuclides & Radiopharmaceuticals*, Sito Web del Centre for Positron Emission Tomography, Austin & Repatriation Medical Centre, Melbourne, Australia, 2000.
[4] B. Bendriem, M. Casey, L. Eriksson, M. Schmand, M. Eriksson, J. Frey, R. Nutt *From PET to PET/SPECT*, Revue de l'ACOMEN, 1999, vol. 5, n$^o$ 2.
[5] D. B. Crosetto *Saving lives through early cancer detection: breaking the current PET efficiency barrier with the 3D-CBS.*, July 30, 2001.

[6] D.B. Crosetto *Comunicato stampa: Sconfiggere il cancro con un'apparecchiatura innovativa per la diagnosi precoce.*, Giugno 2003.

[7] D.B. Crosetto *The 3-D Complete Body Screening (3D-CBS). Features and Implementations.*, IEEE Nuclear Science Symposium and Medical Imaging Conference. Conference Record M7-129, Portland, Oregon, 2003.

[8] D. B. Crosetto *3D-Flow$^{TM}$ DAQ-DSP IBM PC board for Photon Detection in PET and PET/CT.*, IEEE Nuclear Science Symposium and Medical Imaging Conference. Conference Record M3-130, Portland, Oregon, 2003.

[9] J. Franquiz *Positron Emission Tomography: Current Instrumentation, Applications and Radiopharmaceuticals.*, Biomedical Engineering Institute, Florida International University, Miami (2002).

# Resistive Plate Chambers for PET detectors.

Gabriele Sani[1]
Department of Nuclear and Theoretical Physics
- University of Pavia -
via A. Bassi 6 - i27100 PAVIA (PV) - ITALY

[1]Research contract PRIN no. 2003029728 - MIUR

# 1 Introduction

Gas avalanche detectors, conceived and heavily employed in particle physics, have numerous applications in life sciences. In medical diagnostics, gaseous detectors are currently employed in digital X-ray radiography [1] and angiography [2]. High-pressure gas ionization chambers have been used, e.g., X ray sensors in Computerized Tomography (CT). Since they have a rather limited sensitivity to energetic photons, methods were found to couple position-sensitive gas avalanche detectors to solid gamma converters. This is the case in gamma cameras equipped with thick metal-grid converter[3]a. Such devices were routinely applied to medical inspection and, more recently, to a Positron Emission Tomography (PET) mode, for a high-resolution 3D small-animal imaging [3]b. Small-animal PET cameras were also developed, where UV photons from $BaF_2$ crystals are detected in wire chambers operated with a photosensitive (TMAE) gas[4].

There are several other applications of gaseous detectors as diagnostic tools in biomedicine, e.g., beam monitoring in radiotherapy[5], radiation dosimetry, etc. One of the most interesting new fields where their unique characteristic could be exploited is the Positron Emission Tomography (PET).

PET is a radiotracer imaging technique in which tracer compounds labelled with positron emitting radionuclides are injected into the object under investigation. After a short path length[6] the positron annihilates with an electron of the medium emitting simultaneously two 511 KeV photons (almost) back to back. The detection of both photons (whose coupling is made through a check on their time of flight) identifies the occurrence of an annihilation along the chord connecting the detection points. Since the $\beta^+$ emitters are linked to some physiologic substrate such as glucose or oxygen, mapping the density of the positron sources after a while gives a measure of the rate of activity inside the human body. These tracer compounds can thus be used to track biomedical and physiological processes, with applications ranging from the early detection of cancer to neurophysiology studies (see Tab. 1).

| Agent | Used to Monitor |
|---|---|
| $^{18}$F fluorodeoxyglucose | Regional glucose metabolism |
| $^{18}$F sodium fluoride | Bone tumors |
| $^{11}$C methionine | Amino acid uptake/protein synthesis |
| $^{11}$C choline | Cell membrane proliferation |
| $^{11}$C deoxyglucose | Regional brain metabolism |
| $^{15}$O oxygen | Metabolic rate of oxygen utilization/OEF |
| $^{11}$C carbon monoxide | Cerebral blood volume |
| $^{15}$O carbon monoxide | Cerebral blood volume |
| $^{15}$O water | Cerebral blood flow |
| $^{15}$O carbon dioxide (Inhaled) | Cerebral blood flow |
| $^{11}$C butanol | Cerebral blood flow |
| $^{11}$C N-methylspiperone | Dopamine D2 and Serotonin S2 receptors |
| $^{18}$F N-methylspiperone | D2 and S2 receptors |
| $^{11}$C raclopride | D2 receptors |
| $^{18}$F spiperone | D2 receptors |
| $^{76}$Br bromospiperone | D2 receptors |
| $^{11}$C carfentanil | Opiate mu receptors |
| $^{11}$C flumazenil | Benzodiazepine (GABA) receptors |

Table 1: PET agents and their uses.

The development of image reconstruction techniques for the detection of very small fluctuations in density (e.g., cancerous tumors at the early stage of their formation) requires both high-resolution, high-rate detectors and advanced imaging techniques. During the past 30 years many efforts have been made to develop scintillator-based PET detectors and, short of a great technological breakthrough, it seems that their limits have almost been reached. Researche in this field focuses mainly on the growth of crystals with a high light output, a good space and time resolution. However, the possible outcome of this research seems to be limited, while the introduction of gas-based detectors seems to be a more promising alternative path to follow.

In this paper we show a novel feasability study to build a Resistive Plate Chamber (RPC)[7] PET detector with high efficiency, high space resolution and high time resolution. All the results

shown were produced with the Geant simulation toolkit (version 4.7.0), the CLHEP libraries (version 1.8.1.0) and compiled with gcc 3.3.3 under Slackware Linux.

# 2 Scintillator vs RPC: comparison for a PET detector

## 2.1 Efficiency

### 2.1.1 Geometrical efficiency

The PET diagnostic technique offers a unique capability of acquiring methabolic images and has been proven to yield valuable information, in oncology and *in vivo* studies. The major limitation of PET imaging scanners is largely due to the high price of scintillator-based detectors.

One of the most expensive components of such detectors are the scintillator crystals. This is due to the very strict requirements on their characteristics, because they must be blessed with a very high detection efficiency, high light output and good time and energy resolutions. Thus, the major limit in the design of a scintillator-based PET detector is given by the strong dependance of its working parameters on the physical features of every single piece of them. Usually crystals are no more than 25 mm thick in order to have a good spatial resolution (otherwise their parallax and deepness of interaction uncertainties would be too large). They must have also both a good light output to reduce data losses on the coupling to the photomultipliers and a good time and energy resolution to cut noise. Growing such highly specific crystals is still a very expensive process: BGO ($Bi_4Ge_3O_{12}$) crystals 25 mm thick ("commonly" used for PET detectors) cost 100 \$/cm$^2$ odd, while other more exotic crystals can have a much higher price.

Even with such expensive crystals, the resolution of the detector is limited by the parallax effect. To reduce the probability of having photons far from the normal to the crystal surface, PET detectors are commonly built with a cylindrical setup (90 cm diameter for

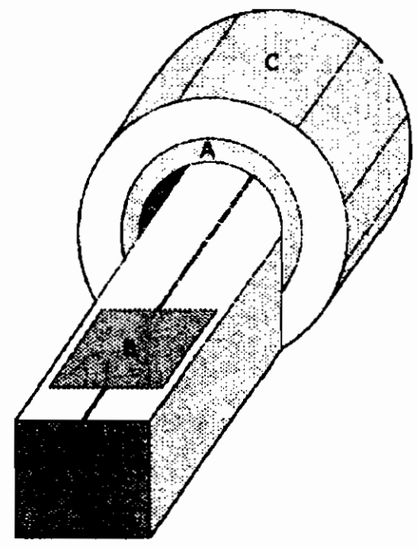
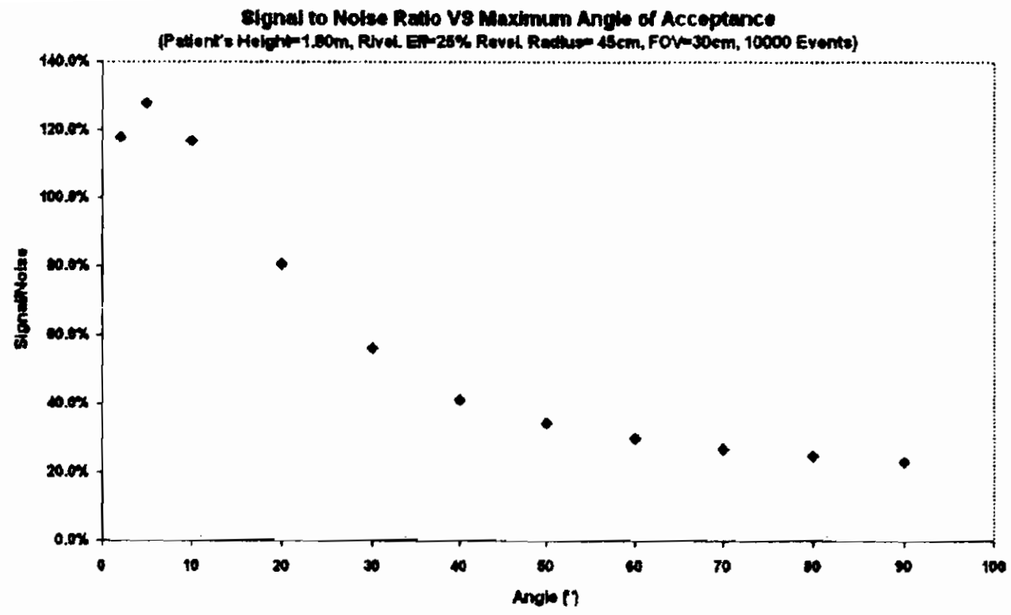



Figure 1: PET detector sketch (elements are to scale): (A) sensible region (B) where the patient lays (C) detector ring.

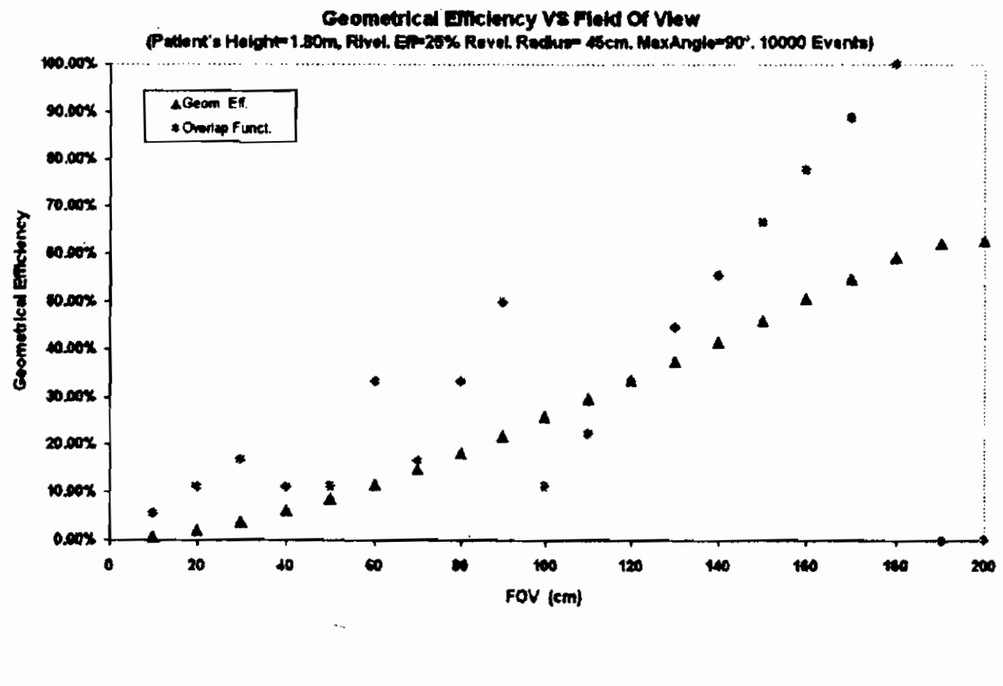
a total-body detector) where the patient is to lay still along the axial line (see fig. 1). To further reduce the parallax incidence the crystals are separated by thin metal layers usually made of lead or tungsten (inter-plane septa) whose depth along the crystals can be modified to tune the rejection efficency of non-perpendicular photons.

The septa can improve image contrast reducing the amount of scattered gammas up to 10-15% of the total counts acquired [9] but, obviously, even on a small detector, 30 cm along thee axis, called Field Of View (FOV) the cut on the maximum acceptance angle reduces greatly the geometrical efficiency $\varepsilon_g$ of the detector. This cut is fundamental since random coincidence events provide incorrect information which degrades the final image. According to the septa depth the scans are commonly said to be either 3D-scans (no septa between crystals) or 2D-scans (fully separated crystal planes). As shown in fig. 2a the signal to noise ratio greatly improves with a harsh cut($< 45^o$) on the angle, but soon it drops when the cut strongly reduces the data. The maximum and the rise in the gradient depend heavily on the FOV.

This detector design has clearly two drawbacks: the limited acceptance angle (a flaw made even harsher by the presence of the inter-plane septa) and the increased cost to cover a wider sensible




a)




b)

Figure 2: a- Signal to noise ratio vs maximum acceptance angle of the incoming photons; b- Geometrical efficiency vs. the FOV.

surface (since the number of crystals needed to cover a cylinder is very high).

The limited acceptance angle, the parallax effect and the high cost to cover large sensible areas strongly limit the geometrical efficiency of total-body PET detectors: they are commonly built with 30 cm FOV. Thus only a very small fraction of the emitted photons can be collected as data while all the others are completely lost or, even worse, contribute to the noise. Building a longer scintillator-based detector is unfeasible since the limited improvement in resolution, due to the limited acceptance angle given by the inter-plane septa and the parallax effect, would not justify the increased cost. Thus it has become common practice to move the detector along

the patient's body to scan it. In order to ensure that there will be no "holes" in the reconstructed image the scans overlap slightly every time. The ratio of the overlap in the lenght is a function of the patient's height. As shown in fig. 2b, while $\varepsilon_g$ is smooth, the overlap function shows abrupt steps due to the change in the number of scans needed to cover the patient's body. Thus to scan the whole body of a patient with such a limited FOV the data acquisition time is multiplied (on average) a sevenfold, since the data have to juxtapose a bit to ensure that no "holes" are left. A longer FOV would greatly improve the geometrical efficiency, increase the signal to noise ratio, shrink the acquisition time, and reduce the tracer dose that has to be injected.

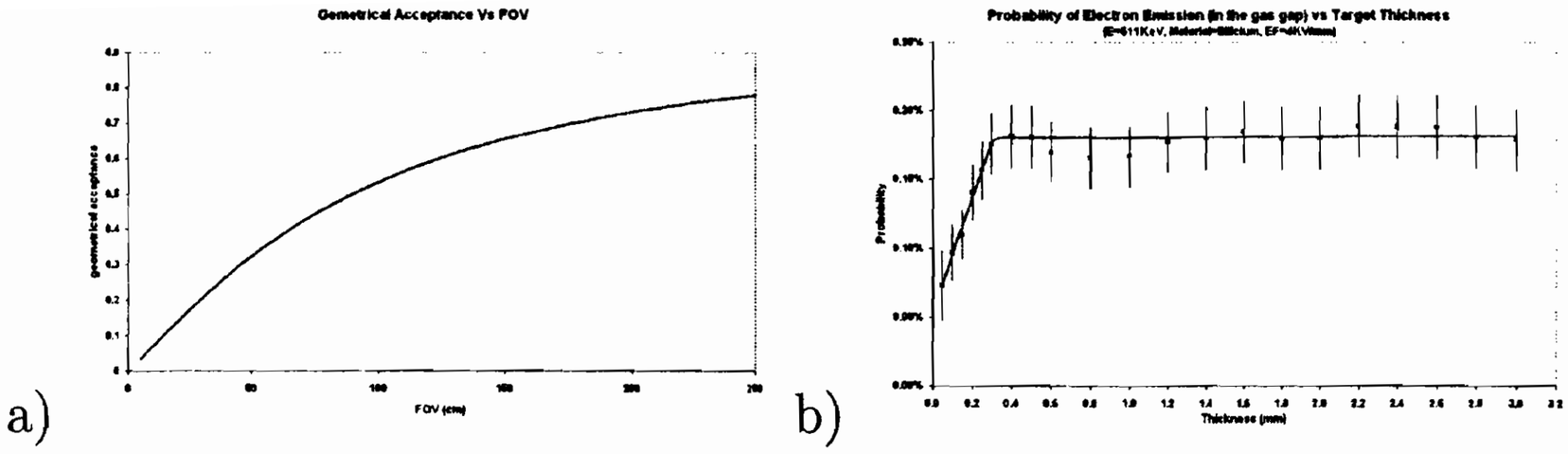
For a source in the middle of the detector the geometrical accaptance ($\psi_{G0}$) of a cylindrical sensitive region is:

$$\psi_{G0} = \frac{2}{\pi} \arctan \frac{FOV}{D} \tag{1}$$

where $D$ is the detector diameter ($FOV$ is simply the detector axial lenght).

For a source not in the middle of the detector the geometical acceptance is given by the above formula where the $FOV$ is to be replaced by the minimum distance between the source and the detector's edge. The total geometrical acceptance ($\psi_G$) is given by the product of the probability of having a source inside the patient's body and the FOV times the probability of having both the photons inside the detector geometry, as shown in fig. 3a, i.e.:

$$\begin{aligned}
\psi_G &= \int_{-\infty}^{\infty} P(x \in \text{body}) \cdot P(x \in FOV) \cdot P(\gamma s \in \text{sensitive surface}) \\
&= \begin{cases} 1/h & \text{if } h_1 < x < h_2 \\ 0 & \text{otherwise} \end{cases} \\
&= \begin{cases} 1/FOV & \text{if } FOV_1 < x < FOV_2 \\ 0 & \text{otherwise} \end{cases} \\
&= \begin{cases} \int_{FOV_1}^{FOV_2} \frac{2}{\pi} \arctan \frac{2\min(t-FOV_1, FOV_2-t)}{d} dt & \text{if } h_1 < x < h_2 \\ 0 & \text{otherwise} \end{cases}
\end{aligned}$$




a) b)

Figure 3: a- $\psi_G$ vs. the FOV; b- Probability of a $e^-$ emission in a gas gap vs the target thickness (with a 4KV/mm electic field applied).

$$= 1/\pi \int_{\max(h_1,FOV_1)}^{\min(h_2,FOV_2)} \frac{2}{\pi} \arctan \frac{2\min(t-FOV_1, FOV_2-t)}{d} dt$$

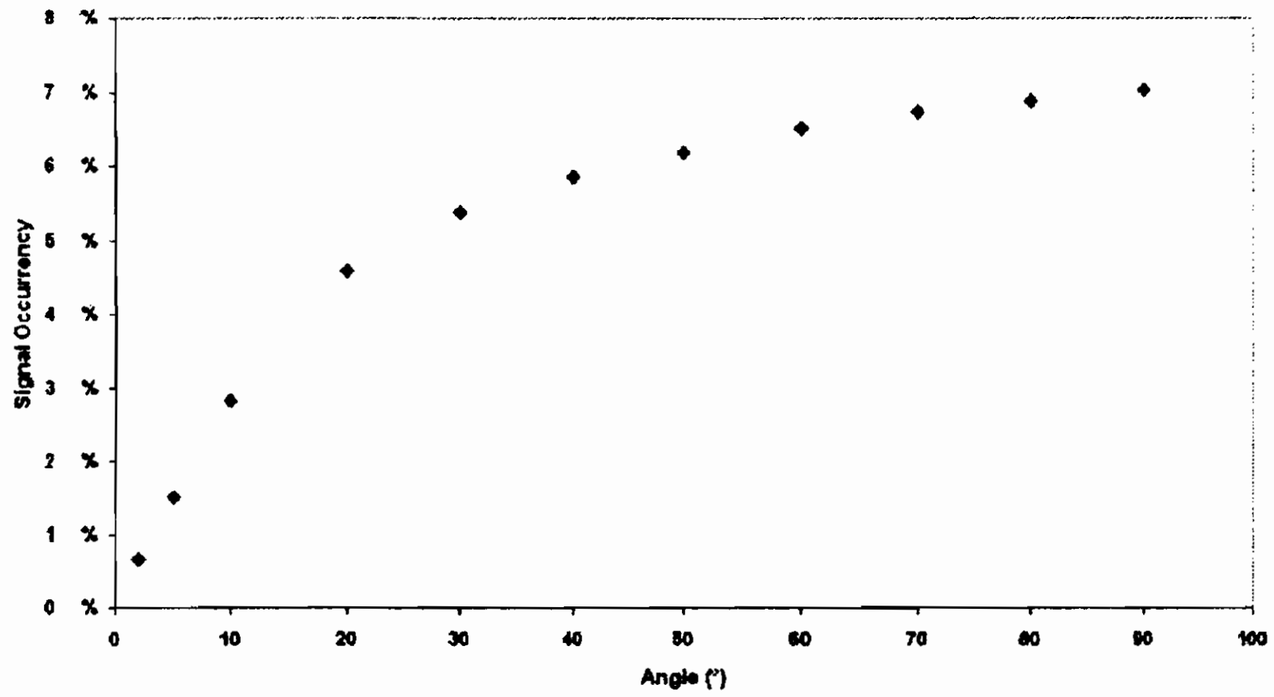
$$\approx 1 - \int_{\min(h_2,FOV_2)}^{\max(h_1,FOV_1)} \frac{2}{\pi} \arctan \frac{\min(x-FOV_1, FOV_2-x)}{D} dx$$

where $h_2 - h_1$ is the patient's height, $FOV = FOV_2 - FOV_1$ and $d$ is the radial distance of the source from the detector surface.

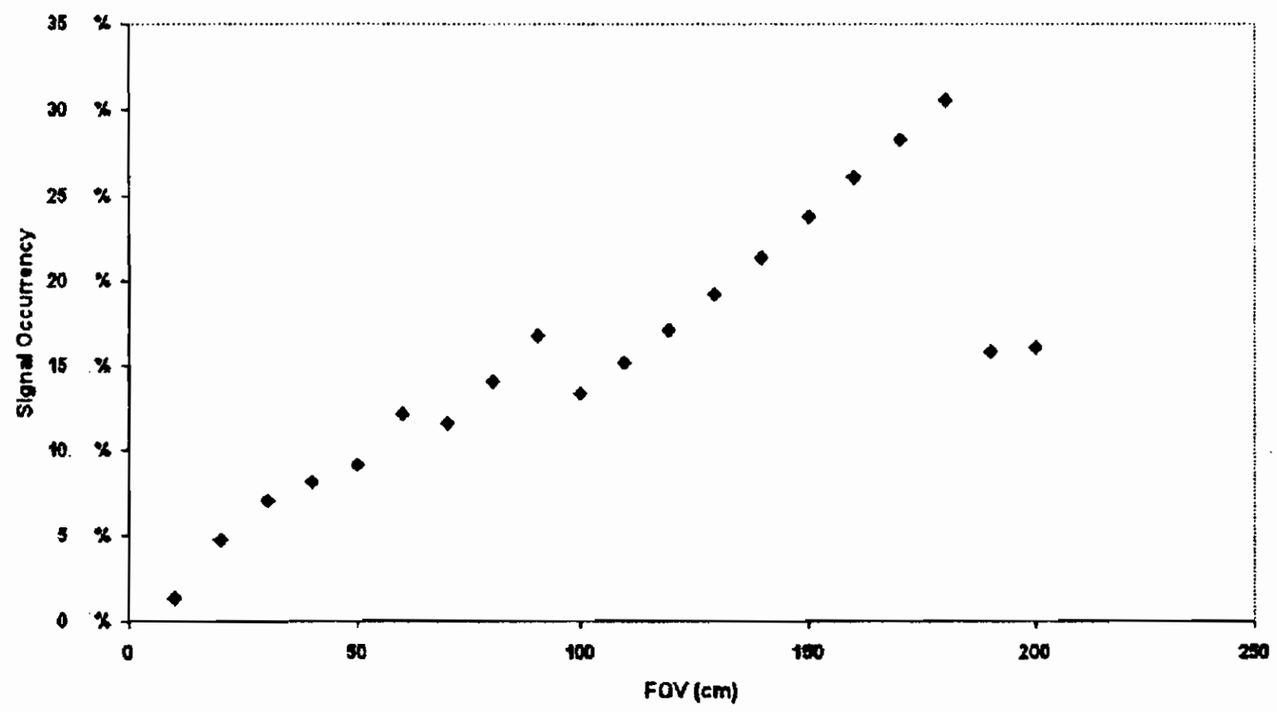
In an RPC,increasing the solid angle coverage would have almost no drawback on the resolution; in fact the "effective thickness" of the converter plate in an RPC is about 300 $\mu$m, a value that depends upon the converter material, while the efficiency depends on the applied high voltage (HV), see fig. 3b). The parallax distortion of the signal can be handled at least partially by the electronic readout system. The low cost of RPCs (near 100 $\$/m^2$: about 10000 times less than a BGO crystal detector) makes it affordable to build a RPC-PET detector with much longer FOV; the use of RPC detectors may allow the assembly of full-body-long FOV scanners. Since the detector response is a function of several parameters it is convenient to study its behaviour through simulations by Monte Carlo techniques. This has been done temporarily assuming that:

1- no photons undergo scattering between the source point and the detector surface;

2- there are as many detectors as needed to scan the whole of the patient's body and to overlap the FOV of every such scan in order to ensure a good efficiency along its full lenght;




a)




b)

Figure 4: a- Signal rate vs maximum acceptance angle; b- Signal rate vs FOV.

3- no time cut is used since it would simply reduce the noise occurrency (thus the signal to noise ratio would accordingly rise).

The results of these simulations are shown in fig.s 2, 4 and 5.

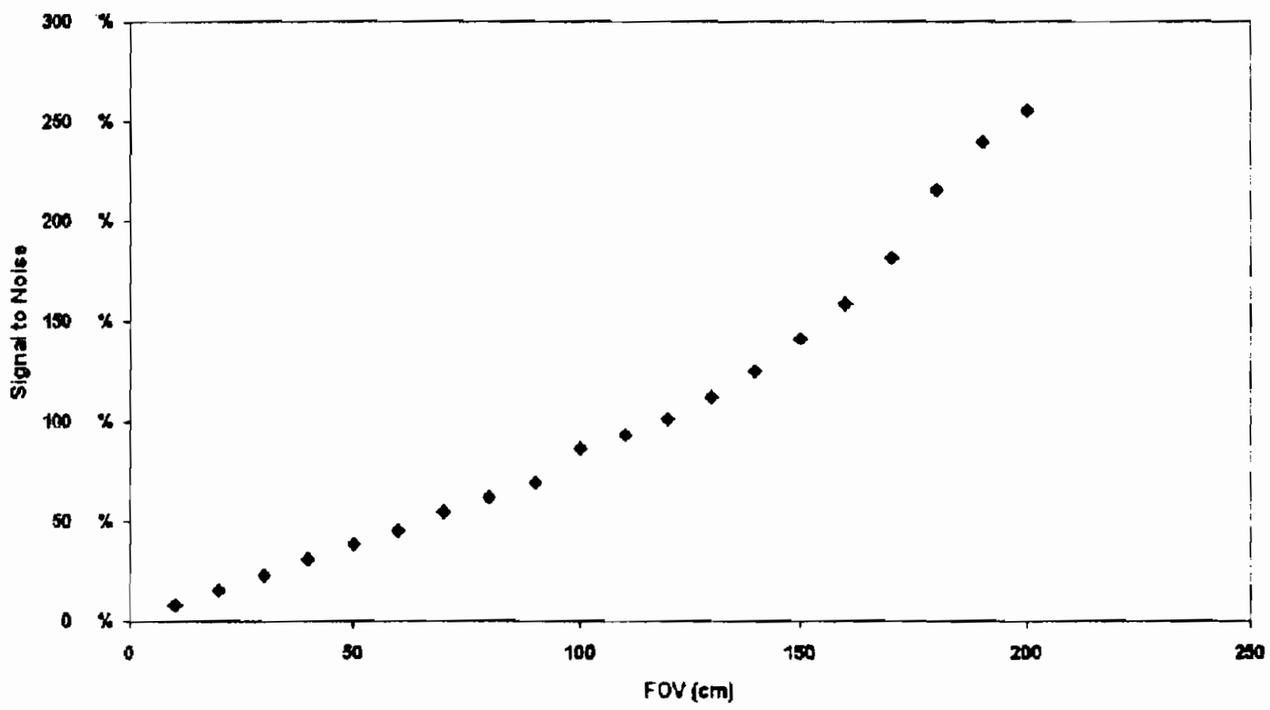
A FOV as long as the patient's body would greatly increase the data set and reduce the noise since photons emitted outside the detector length may enter the sensible area and become spurious data. The required data acquisition time would be shorter, greatly

reducing the minimum dose to be injected in the patient's body.

### 2.1.2 Scattering amplitude

As long as only the detection of photons is of interest, an RPC can be modeled as a converter plate plus an electron selector plus a signal amplifier.

a)




b)

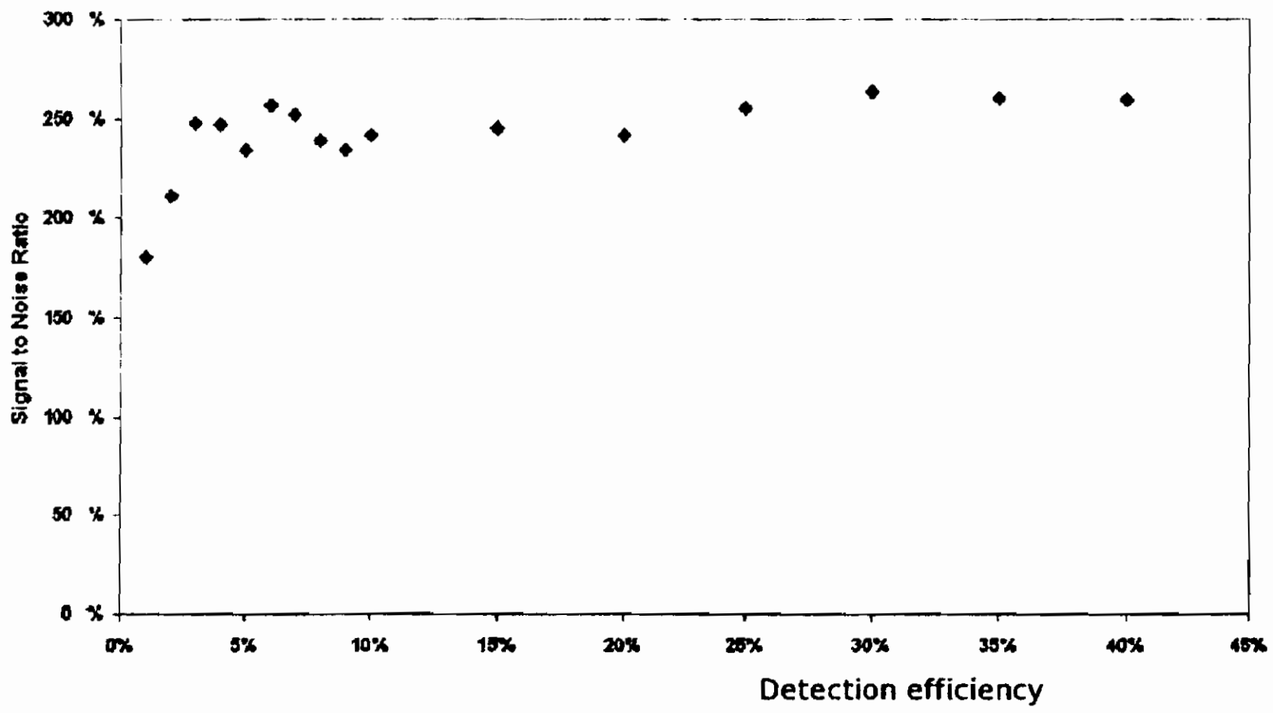



Figure 5: a- Signal to noise ratio vs FOV; b- Signal to noise ratio vs detector efficiency.

The detection efficiency is ruled mainly by the conversion efficiency given by the (photoelectric emissions and) Compton scattering cross section shown in fig. 6.

High Z materials have a rather large gamma conversion probability (see App. 4.2) thus it would be favourable to employ high Z target layers. Multiple electron scattering inside the converter plate could lead to electron capture, thus the probability of recording a signal depends on the distance the electron has to cross to reach the gas gap where the signal amplification starts (see fig. 3b).

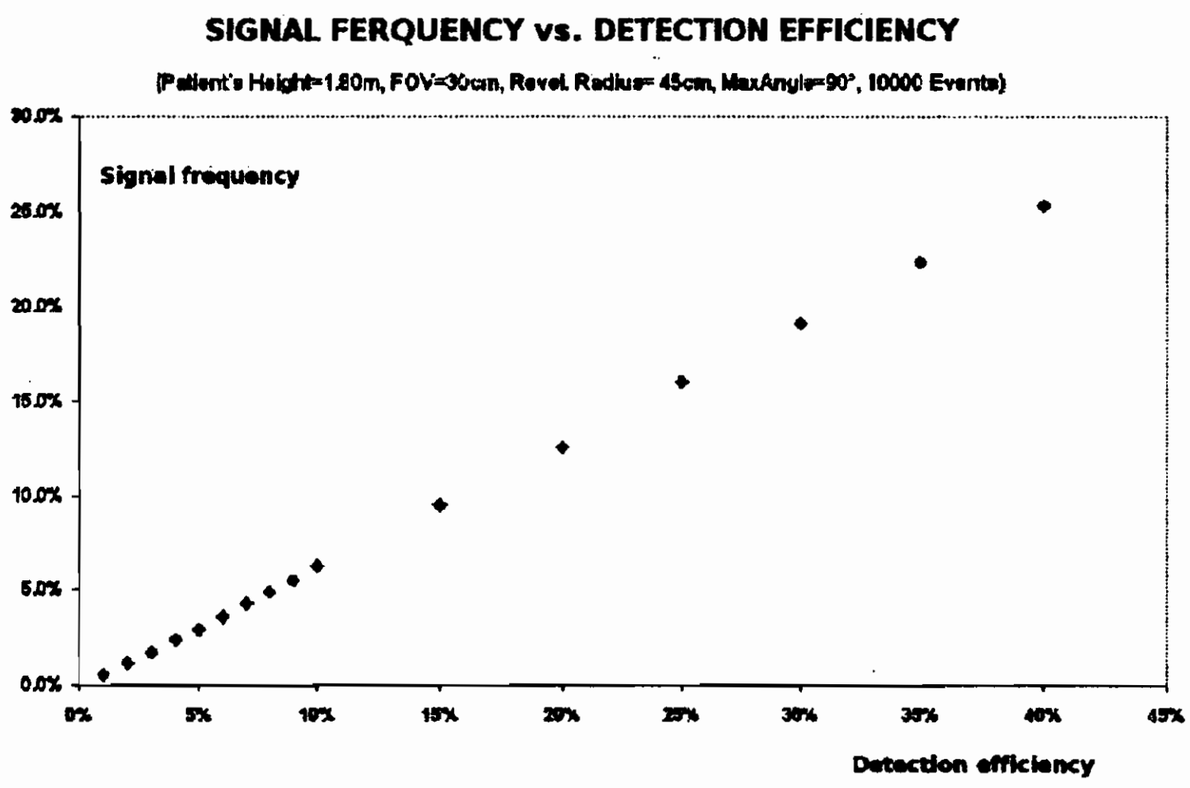



Figure 6: Total (Compton + Photoelectic) scattering probability inside a silicon target.

A strong electric field between the electrodes increases the emission probability as the electrons are "pushed" toward the gas gap.

In ordinary PET scanners, in order to build detectors and improve the resolution much efforts have been made to grow crystals with the suitable scintillating characteristics *and* high Z. In RPC this problem can be circumvented, since there are much less strict requirements on the materials to be used. For example, resistive plates can be built with a high Z material, or, since there is no need for an homegeneous material, they can be coated with a thin high Z paint or drugged near the surface. Even the detector layout can

be tuned: multiple high Z target planes can be inserted in the gas gap. Adopting a MultiGap layout (MRPC) is an effective way to increase the photon detection efficiency.

### 2.1.3 Data acquisition

Data acquisition is also a crucial feature of a detector since it rules the maximum acquisition rate, the probability of detecting impinging photons and the signal-to-noise ratio.

Scintillating crystals are optically coupled to photomultipliers (PMT), which collect the light output and convert them into electric signals. This coupling has a far from optimum efficiency, since less than 30% of the signal is able to pass through the light interface and become data to be processed. In RPCs every electron, as soon as it reaches the surface of the gas gap, undergoes a very strong acceleration that produces an avalanche between the electrodes (the gain is about $10^8$). The produced ion and electron avalanche yields a fairly strong electromagnetic signal that induces a pulse into the readout electrodes. This signal is read out into the data acquisition modules for analisys. Therefore *all* photoelectrons that reach the surface of the gas gap and overcome a small threshold against iontrinsic electronic noise are recorded.

The spatial resolution in scintillator-PM detectors is given mainly by three factors: the spread of light inside the crystal, the resoluton of the photomultipliers and the optical features of their interface. On the contrary, in RPCs the interaction point is quite near (hundreds of $\mu$m) to the emission point in the gas gap and even the avalanche has got a very small extension. A good estimate of this factor can be given using the distance between the electrodes ($d$), the intensity of applied HV ($V$):

$$\begin{cases} C = \frac{\varepsilon A}{d} \\ C = \frac{Q_{TOT}}{V} \end{cases} \Rightarrow A = \frac{dQ_{TOT}}{\varepsilon V} \qquad (2)$$

where $\varepsilon = \varepsilon_r \varepsilon_0$ is the dielectric constant of the RPC material while $Q_{TOT}$ is the total produced charge that must be absorbed to recover the initial state. With the typical working parameters ($V/d = 5000$

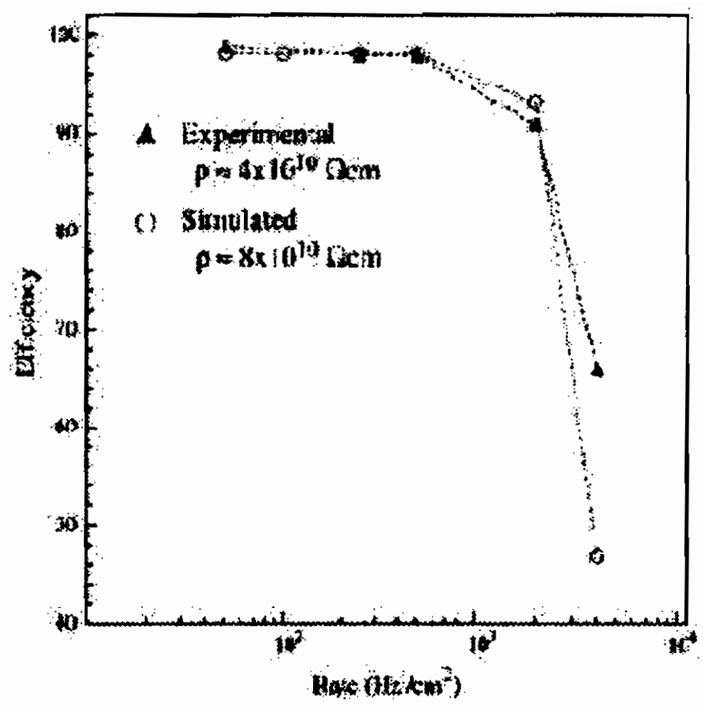



Figure 7: The counting rate capability of RPCs under continuous irradiation [11].

V/mm, $\varepsilon_r = 5$, $Q_{TOT} = 10$ pC the area affected is about 0.04 mm$^2$. Assuming a "drop-shaped" development of the avalanche its radius would be just 80 $\mu$m.

### 2.1.4 Maximum event rate capability

The maximum event rate is limited mainly by the unsensitive area and the deadtime of the individual detector modules but can be limited also by the speed of the coincidence and address logic electronics. In recent literature it has been stressed that there is a critical need in PET for a scintillator with a decay time significantly shorter than the 300 ns of BGO, such as LSO or $PbSO_4$ [8], [10].

It should be asked whether these counters might withstand the counting rates desiderable for PET. The counting rate capability of glass MRPCs under continuous irradiation is shown in fig. 7.

Under standard operating conditions the maximum rate can be as large as about 1-2 kHz/cm$^2$.

Since the maximum dose injected (used for the total-body PET) is about 10 mCi, the total maximum emission rate ($M_{ER}$) on the full solid angle is about 370 MHz. The typical dimension of the

detector rings (whose diameter is about 90 cm for a total-body PET) and the dose injected yields a statistical upper bound on the emission rate on the surface ($R_S$) of the detector:

$$\begin{aligned} R_S(a+\delta a, b+\delta b) &= M_{ER}/(\delta a * \delta b) \\ &\leq M_{ER}/(\pi D^2) \\ &\leq 15kHz/cm^2 \end{aligned}$$

Where $(a, a+\delta a)$ and $(b, b+\delta b)$ define a unit of surface on the detector and the approximation is given for a *spherical* sensible area with diameter $D$. Since every converter plate has got about 0.5% efficiency, every RPC can have as much as $600/(1.5 \cdot 10^4 \cdot 0.5\%) = 8$ target layers without significant loss in efficiency due to their rate capability. This rough estimate is bound to be very conservative, since we worked on the rather naive assumption that *all* the radiotracer was deposited in the same place. The half life of the radioisotopes commonly used for Total Body PET scans ($^{18}$F, whose life time is 109 min) is of the same order of the time spent after the injection to let the chemical cocktail diffuse in the body (about 20 min). Thus the maximum number of possibly "effective" layers is even higher, more than the feasible number of layers to be inserted in a Multigap RPC. Therefore it seems that the maximum system counting rate will not be limited by the RPCs rate capability.

### 2.1.5 Deadtime data loss.

Every measurement system exhibits a characteristic deadtime[12]: since the pulses produced by a radiation detector have a finite time duration, when a second pulse occurs before the first has disappeared, the two pulses will overlap to form a single distorted pulse [13], [14].

Depending upon the system, one or even both particle arrivals will go unrecorded. This loss of counts changes the measurement statistics and the statistical moments. The statistics of the coincidence-counting process for detectors affected by deadtime is of fundamental importance to the problem of statistical image reconstruction, as elaborated in [15] for single-photon counting systems. Counting

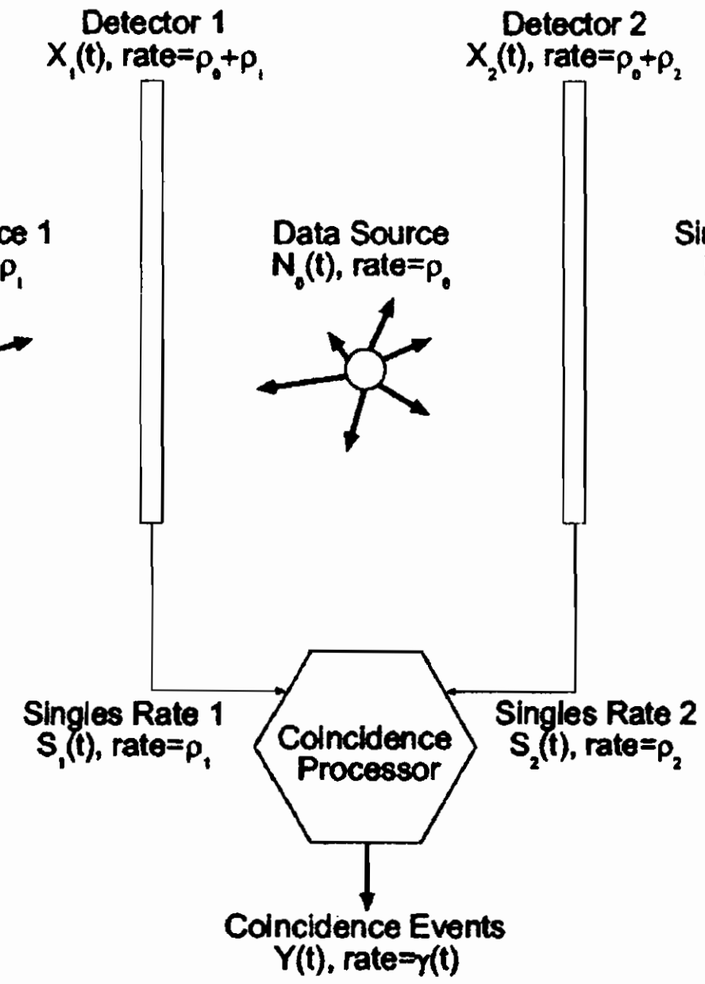



Figure 8: Model for coincidence counting.

systems are often characterized as either non-paralyzable or paralyzable (extendable). In a paralyzable system, each particle arrival, whether recorded or not, produces a deadtime of length $\tau$; if there is an arrival at time $t$ then any arrival from $t$ to $t + \tau$ will go unrecorded.

This section investigates the and variance of the coincidence-counting processes, for accidental coincidences and for total coincidences (both accidental and genuine), also known as prompt coincidences in the PET literature.

In the scenario illustrated in fig. 8, the random process $N_0(t)$ counts the number of particle pairs emitted simultaneously from the "genuine-coincidence source" and arriving at the two detectors in the time interval $[0, t]$. It can be stated that the particles in each $N_0$ pair arrive simultaneously at the two detectors (similar results can be obtained even for non-simultaneous arrivals). The random processes $N_1(t)$ and $N_2(t)$ count particles originating from the random singles sources as they arrive at the two detectors. It can be

assumed also that $N_1(t)$ and $N_2(t)$ are stochastically independent. To preclude the possibility of multiple simultaneous arrivals, it can be assumed that any arrival process $N_i(t)$ considered in this paper has the property $P[\Delta N_i(t) \geq 1] = 0$ for any time $t$; where $\Delta N_i(t)$ denotes the increment in $N_i(t)$ at time $t$. Let $Y(t)$ be the sum of three factors:

$$Y(t) = Y_0(t) + Y_1(t) + Y_2(t) \tag{3}$$

i.e. $Y_0(t)$, the number of recorded coincidence events corresponding to simultaneous arrivals, at time t, at the two detectors; $Y_1(t)$ and $Y_2(t)$ be respectively the number of recorded coincidence events during $(0, t]$ that have singles event arriving late at detectors 1 or 2; let also $\Gamma$ be the istantaneuous rate and $\rho_1$, $\rho_2$ the counting rates of, respectively, $N_1(t)$ and $N_2(t)$. Then, at the time $t$:

$$\Gamma(t) = \lim_{\delta \to 0} \frac{E[Y(t+\delta) - Y(t)]}{\delta} \tag{4}$$

so we can write the mean value of accidental coincidences as:

$$\begin{aligned} E[Y(t)] &= \int_0^t [\Gamma_1(s) \int_{s-r}^{s} \Gamma_2(u)du + \Gamma_2(s) \int_{s-r}^{s} \Gamma_1(u)du]ds \\ E[Y(t)] &= \Gamma t \end{aligned}$$

while the variance is:

$$E[Y(t)^2] = \Gamma t(1 - \Gamma\tau + (2 - \tau/t)) + \frac{\Gamma^2}{24}(5r^2 - 16r(t - \tau)) \tag{5}$$

where $\Gamma = 2r\Gamma_1\Gamma_2 = 2r\rho_1\rho_2 \exp^{-(\rho_1+\rho_2)\tau}$.

It is worth comparing the coincidence-counting process to the singles-counting process. Deadtime causes the singles-counting process to be significantly non-Poissonian, in the sense that the variance of the process is significantly less than the mean value. In contrast, the variance of the coincidence-counting process is very close to the mean so it can be treated as poissonian. In PET, accidental coincidence (AC) events occur when two photons originated in separate annihilations are detected within the coincidence time window. The effect of AC events is most severe for sourcess

with low true coincidence rates, such as those traversing the thorax. Most PET scans are compensated for AC events by real-time subtraction of delayed-window coincidences. This precorrection yields the proper ensemble mean, but destroys the Poisson measurement statistics. To preserve Poisson statistics, one should acquire separate transmission and random data. However, because of hardware and data storage space limitations, most PET centers currently collect and store only the random precorrected data. Scintillating crystals have a deadtime in the order of 300 ns while RPC have a much better time response (less then few ns, down to 50 ps) thus the statistical occurence of AC events is much smaller. The deadtime in an RPC is the time needed to recover the initial condition: the HV between the electrodes is reduced by the charge contained in the avalanche impinging on the resistive plates, then the detector is charged again by the external power supply with an exponential behaviour characterized by an "electrode time constant", $\tau_0$, given by the formula:

$$\tau_0 = 2\rho_r\varepsilon_0(2\varepsilon_r + \frac{b_r}{g}) \tag{6}$$

where $\rho_r$, $\varepsilon_r$ and $b_r$ are, respectively, the bulk resistivity, relative dielectric constant and thickness of the resistive plates[7]. The "electrode time constant" is the leading factor when the deadtime is to be calculated.

### 2.1.6 Non-uniformity of data sampling

Reconstruction algorithms usually are exployted with the assumption that data is sampled on a continuous space and that the sensible region is uniformly efficient[7]. Since the physical signal is brought to the electronics through a discrete set of elements, this assumptions are to be questioned, and could lead to some degree of signal distortion or loss in efficiency. On top of this, real imaging systems are subject to a number of physical effects that make their response space-variant and image-dependent.

In PET, nonuniform sampling and crystal penetration effects also lead to space-variance[17]. There are also object-dependent

attenuation effects that can alter the resolution. When all these effects are not compensated, the reconstructed images can suffer from quantitative inaccuracies and geometric distortions due to the anisotropic response. These nonuniform resolution properties can complicate the image reconstruction since image-dependent resolution properties make the comparison of different images or different features in the same image more difficult. For tomographic applications, there is a number of analytical methods that can compensate for space-variant physical effects. Till now, a number of methods for PET with irregular sampling functions have been developed [18]. Under proper conditions, some techniques can yield images free of resolution nonuniformities, however usually noise models are ignored. An alternative approach is to perform Maximum Likelihood (ML) reconstruction using an accurate system model. When the system model incorporates the system geometry and all the physical effects, it is often possible to obtain image estimates that have nearly "perfect" resolution when a pixelated object model is used. This means that the local impulse response defined in [19] is a Kronecker impulse. Such estimates require iterating the algorithms used to maximize the ML objective until convergence. Unfortunately, such images usually appear overly noisy due to the ill-conditioned nature of inverse problems. A number of solutions have been proposed to improve the appearance of such images: a- The Maximum Likelihood images can simply be post-filtered with a shift-invariant blur. If the ML image has "perfect" resolution, then the blurred image will have uniform resolution. However this still requires a fully converging solution to an unregularized problem, which may take very many iterations.

b- Iteration can begin with a uniform image and can be stopped prematurely yielding a smoother result. However such images will have nonuniform resolution properties [20]-[22].

c- The problem can be regularized, improving convergence rates and image quality. While one can use sieves [23] to regularize the problem, the appropriate kernels may not always exist for a desired resolution and system model.

Additionally, a space-invariant filter cannot provide uniform res-

olution for space-variant systems. Another form of regularization is the penalized-likelihood approach, where a penalty term is added to the objective function that discourages rough images. Standard space-invariant penalties yield nonuniform resolution properties even for space-invariant systems due to the implicit data weighting of the penalized-likelihood objective [23]. Space-variant penalties have been developed that yield uniform resolution properties [24], [25]. In principle, these techniques may be applied to a wide range of space-variant imaging systems to correct for both the implicit data weighting and to compensate for the various physical effects that make the system space-variant. Such methods would provide for easy resolution control, where one needs only to specify the desired point spread function. However, these space-variant penalties are data dependent, and they must be computed for each data acquisition. Moreover, fast techniques for calculation of the penalty term have only been developed for space-invariant systems [24]. The main obstacle to the application of these techniques to space-variant systems, in general, remains the efficient calculation of the penalty.

## 2.2 Theoretical limits on resolution

To build a good theoretical framework in which to sketch the project of an improved PET detector, it is necessary to underline the intrinsic limits on data resolution. The aim of a PET image reconstruction is to pinpoint tumors and their methabolical information. This analisys is built on a few assumptions:

1- the density of the positron sources is a good marker;

2- the map of the positron sources is faithfully reproduced by the map of the annihilation points;

3- the emitted photons carry enought information to reconstruct the density of annihilation points;

4- the reconstructed image is correct;

| Isotope | $^{18}$F | $^{11}$C | $^{68}$Ga | $^{82}$Rb |
|---|---|---|---|---|
| Max $\beta^+$ energy (MeV) | 0.64 | 0.96 | 1.90 | 3.35 |
| fwhm (mm) | 0.13 | 0.13 | 0.31 | 0.42 |
| fw(0.1)m (mm) | 0.38 | 0.39 | 1.6 | 1.9 |

Table 2: Positron range factors for four isotopes in water [28].

### 2.2.1 Field of application of PET imaging techniques

Recent experiments have shown that radiotracer compounds diffuse through the whole body and their rate of deposition is a valuable diagnostic tool. Anyway, it cannot be used to track every biomedical and physiological processe nor to diagnose every kind of tumor. Methabolical data must be well know and must taken into account to avoid misinterpretation of the reconstructed image.

### 2.2.2 Positron range

The annihilation points do not coincide with the emission points since positrons travel a short path before being annihilated. Measurements of positron range have shown that this distribution consists of a central spike (fwhm $< 0.5$ mm) plus tails that can extend outward for up to several mm [26],[27],[6], [28],[29].

Tab. 2 collect values of full width at half maximum (fwhm), full width at 0.1m [fw(0.1)m], as well as the maximum $\beta^+$ energy for four of the most common positron emitters. The fwhm and fw(0.1)m describe the narrow central region of the distribution, but not the tails. The fw(0.1) reflects the statistical broadening of the entire range distribution. The statistical distribution of end points of positrons emitted by $^{18}F$ sources is shown in fig. 9

### 2.2.3 Noncollinearity

Since the positrons do not come to a complete rest before annihilation, the two 511 keV annihilation photons are not emitted in

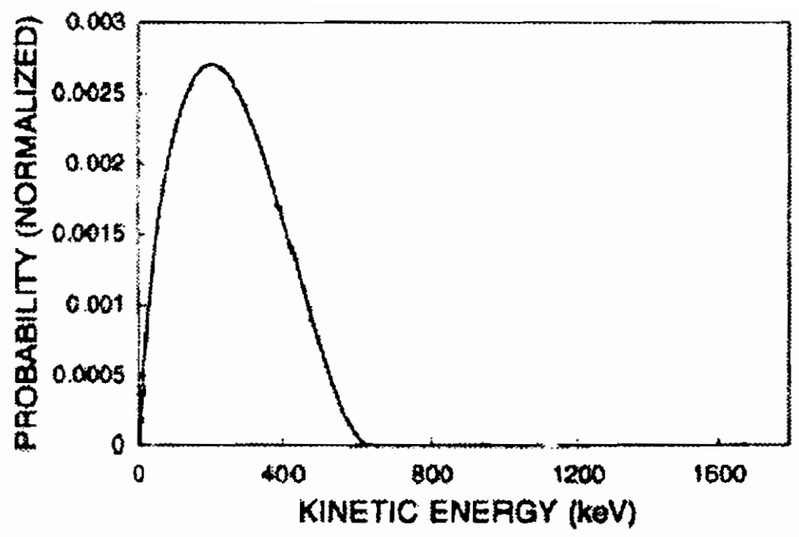



Figure 9: Statistical distribution of $^{18}$F-emitted positron end points in water[30].

exactly opposite directions[31]. The result is a Gaussian angular distribution with about $0.5^o$ fwhm. For a source placed at the center of the tomograph, this introduces a contribution to the resolution of about $0.0022D$, where $D$ is the detector ring diameter (in millimeters). This means that for a total body PET, where $D$ =90 cm, the blurring of the signal is about 2 mm or that the source point is about 1 mm off the line adjoining the two impact points.

The non-collinearity of the photons may have a major influence on the geometry development of a PET detector. Since this blurring is proportional to the distance between the annihilation point and the detector surface, increasing the FOV will have the drawback that the "faraway" data will have a poorer resolution. Changing the detector geometry (eg: using a box-shaped detector with tunable lenghts) in order to reduce the path of flight would reduce the data spreading.

### 2.2.4 Combined formula for reconstructed image resolution

The combined reconstructed image resolution $\Re$, as influenced by detector size $d$, noncollinearity (through the detector array diameter

$D$), the effective positron range $r$ is given by the formula:

$$\Gamma = 1.25\sqrt{(d/2)^2 + (0.00022D)^2 + r^2 + b^2} \tag{7}$$

where an additional factor $b$ is introduced to account for other distortions, such as those given by the parallax effect, the non-uniformity of the data acquisition, etc.

## 2.3 Noise

Noise estimators are perhaps the most complex piece of a simulation. This is due to the randomness of their nature and the difficulty even to pinpoint their statistical features. Nonetheless they are among the most useful tools to determine if a simulation is good enough to fit the data and to test the image reconstruction algorithms. This feature is extremely important for PET image reconstruction techniques, since it has been shown that iterative algorithms and maximum likelihood estimators can be much slower or even fail a reconstruction (mainly they create "noise artifacts") due to the occurrency of "noisy" data. Several techniques have been develped to reach a better signal-to-noise ratio.

## 2.4 Time window

Setting a time window is one of the most common and crucial technicniques to discriminate signals. Events detected within a *preset* time are assumed to be fathered by the same process (occurring along the line joining the two detectors which recorded the events), while events outside the time windows are assumed to be noisy events or singles (the other photon being not detected) and are rejected (see the scketch in fig.10). The coincidence time window used for event detection is based upon the type of scanner. BGO cameras use a coincidence window of 12 ns, for GSO and NaI systems the width is reduced to 8 ns while LSO systems are commonly operated with just 6 ns. Many efforts were spent to grow scintillating crystals with a better time resolution since, obviously, a shorter time window gives a a better noise cut-off. RPCs, due to their superior

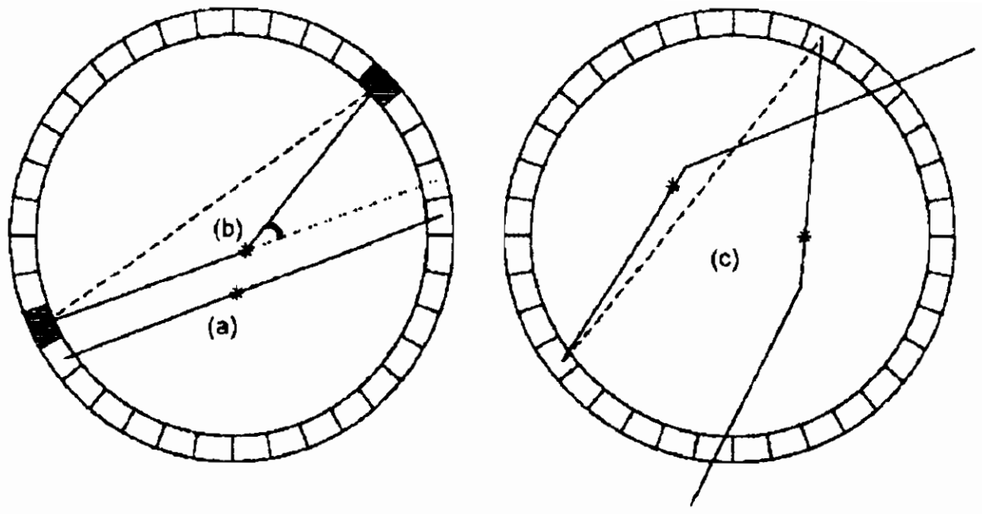



Figure 10: Signals and noise: (a) both photons reach the sensible region unscattered and are recorded, giving a "good" signal, (b) One (or both) photon(s) scatter(s) before reaching the detector, giving a distorted signal, (c) Two photons from two different sources are recorded inside the time window, this is clearly noise.

time resolution, can work with a much shorter time window. It can be so short that they open the path to use the time of flight information to give a rough estimate of the point of interaction *along* the chord. In PET measurements, accidental coincidence (AC) events are a primary source of background noise[33]. AC events occur when photons arising from separate annihilations are mistakenly registered as having been generated by the same annihilation. In order to estimate their occurrence the "singles" method[34] can be employed. However this approach is not widely used because of the necessity to add hardware. Moreover, the rate of the singles usually varies during the data acquisition time[35]. In most PET scans, AC rates are estimated using delayed-window coincidences and the data are precorrected for AC events by real-time subtraction. However real-time subtraction of delayed window coincidences provides an average compensation for AC events but distorts the Poisson statistic [36]. To avoid this problem, one needs to maintain transmission and random delayed coincidences measurements as two separate data sample [37] [38]. The drawback is that collecting randoms data sets separately would double the storage space for the acquired data. That's why in practice most PET centers

collect and archive only the randoms precorrected data.

## 2.5 Energy Cuts and Time Of Flight cuts

Photons crossing a human body may scatter and deviate from their previous path. Since, clearly, the scattering destroys the spatial information about the source carried by the two photons, the reconstruction is impossible, or even worse, misleading. Therefore this events must be regarded as noise and rejected. To accomplish this, scintillator-based PET detectors use an energy cut (usually it is about 20%) to select 0.511 MeV gamma rays. Since RPC detect the photons through two information-disruptive steps (gamma-electron conversion and electron multiplication) this kind of cuts is not feasible in an RPC.

The converter plate physics, however, grants a kind of "built-in" energy cut: even if the scattering amplitude is larger for low energy photons, the mean path that a gamma-produced electron can travel before being re-absorbed is shorter, thus RPCs have a smaller detection efficiency for low energy gammas.

A very interesting way to cut spurious data is to use the Time Of Flight (TOF). The very good timing of Multigap RPCs (up to 50 ps) opens the possibility to use the difference between the time of flights of the two gammas to localize the annihilation point along the line joining the two coincident detectors. The TOF cut has several advantages (see [39]-[48] for further details):

a- the image can be reconstructed with less statistical noise;

b- each annihilation can be placed near the image plane where it occurred;

c- angles can be grouped, which reduces the task of data storage and tomographic reconstruction;

A time of flight PET detector would open new techniques for the data reconstruction; togethere with a much smaller noise ratio given by the enlarged field of view it is very reasonable to think that the image reconstruction would be *significantly* improved.

# 3 Conclusions.

The comparison of the features of a scintillator-based PET detector to those of an RPC-based strongly suggests that the latter could be a far superior PET detector. Clearly the most critical point of the converter plate approach to photon detection is the low quantum efficiency (about $0,5\%$ for every layer, but obviously it depends on the converter plate material). Such a low value is due to the scattering amplitude of the interaction and to the low energy spectrum of the emitted electrons, which limit the maximum "meaningful" thickness of the converter plate. It is common practice to use several layers of RPCs to increase the detection efficiency, since the gamma beam is only slightly depleted crossing a detector.

Montecarlo results suggest that electrons produced by 0.511 MeV gamma rays have a reasonable probability of crossing 200 $\mu$m odd before being re-absorbed. Thus interactions farther than 200 $\mu$m from the edge of the converter plate probably will go undetected and must be regarded as lost data. This feature greatly influences the design of every converter-plate-based PET detector since it determines a strong link between the maximum efficiency (data flux reduction) and the mechanical accuracy of the detector.

## 3.1 Main features of the proposed detector.

Extensive testing and further developments are bound to be done. We have been able to assess several key features of an economic and accurate RPC detector to be employed in the PET medical imaging technique: a Multigap RPC would be a promising choice. The MultiGap design would allow to build a very compact and efficient detector with several favourable features. Inserting, e.g., 6 target layers inside the gas gap the efficiency would be boosted at least a sixfold (the layers could be made of high Z material, so the efficiency would be more than just a sixfold), while the gas gap would be only 2.6 mm (assuming 200 $\mu$m thick layers spaced by 200 $\mu$m). The parallax effect would increase, but the efficiency gain would pay off the small distortion of the signal. Multigap RPC can be made into

very fast detectors with a time resolution of about 50 ps, thus the spurious noise incidence would be minimal (due to the short time window) and the difference between the time of flight could be used not only to cut AC data but also to improve the reconstruction of the source points. The uncertainty in difference of the TOF would be 100 ps, thus the source could be reconstructed both along the adjoining chord and within 3 cm along it. A Multigap RPC PET detector would be much cheaper than a normal RPC one, too: the increased efficiency of the detector would greatly reduce the number of components to be placed in the tower, thus reducing the cost of the electronics and handiwork needed.

# 4 Appendix

## 4.1 A: $\beta^+$ absorbtion

For the majority of $\beta$ spectra, the range curve has a near exponential shape. The coulomb forces that constitute the major mechanism of energy loss for both $\beta$ and heavy charged particles are present for either positive or negative charges. The impulse and energy transfer for particles of equal mass are almost the same, therefore the positron tracks in an absorber are similar to those of electrons and their specific energy loss and range are about the same for equal initial energies. The distribution of the end point of the positron trajectories determines their contribution to the spatial resolution in a PET detector. This distribution is isotropic thus it can be modelled with a one-dimensional function:

$$P(x) = Ce^{-k_1 x} + (1 - C)e^{-k_2 x} \quad x \geq 0 \tag{8}$$

with parameters $C$, $k_1$, $k_2$ that depend on the $\beta$ energy spectra and on the material composing the absorber. So far, they have been measured only in water; the numerical values are shown in Tab. 3

| | $^{18}$F | $^{11}$C | $^{13}C$ | $^{15}$O |
|---|---|---|---|---|
| $C$ | 0.516 | 0.488 | 0.426 | 0.379 |
| $k_1$ ($mm^{-1}$) | 37.9 | 23.8 | 20.2 | 18.1 |
| $k_2$ ($mm^{-1}$) | 3.1 | 1.8 | 1.4 | 0.9 |

Table 3: Coefficents of eq. (8) in water.

## 4.2 B: Photon interaction

### 4.2.1 Photoelectric absorbtion

In the photoelectric absorbtion process, a photon undergoes an interaction with an atom that results in the emission of an energetic electron (sometimes called *photoelectron*) from one of the bound shells. The photoelectron appears with an energy given by

$$E_{e^-} = E_\gamma - E_b \tag{9}$$

where $E_{e^-}$ is the emitted electron energy, $E_\gamma$ is the incoming gamma energy and $E_b$ is the binding energy of the electron in its original shell. A rough approximation of the probability of photoelectric absorbtion per atom over all the energy range is given by:

$$\tau = \text{constant} \times \frac{Z^n}{E_\gamma^{3.5}} \tag{10}$$

where the exponent $n$ varies between 3 and 5 in the region of interest. Since the binding energy is much smaller than the energy emitted in a $e^+ - e^-$ annihilation its contribution can be neglected, thus the energy spectra of photoelectrons is just a peak at (almost) the same energy of the absorbed photon.

### 4.2.2 Compton Scattering

Compton scattering takes place between the incident photon and an electron in the absorbing material. It is most often the predominant interaction mechanism for $\gamma$ energies typical of radioisotope

sources. It is the most dominant interaction mechanism in tissue. In Compton scattering, the incoming photon is deflected through an angle from its original direction. The photon transfers a portion of its energy to the recoiling electron (assumed to be initially at rest), which is then known as a Compton electron. The energies of the scattered photon $h\nu'$ and of the Compton electron $E_e$, are given by

$$h\nu' = \frac{h\nu}{1 + \rho(1 - \cos\vartheta)} \tag{11}$$

$$E_e = h\nu - h\nu' - E_b \approx h\nu - h\nu' \tag{12}$$

where $\rho$ is the energy of the incoming photon in units of energy of an electron at rest. The angular distribution of scattered gamma rays is predicted by the Klein-Nishina formula for the differential scattering cross section:

$$\frac{d\sigma}{d\Omega} = \frac{Zr_0^2}{2} \frac{1 + \cos\vartheta}{[1 + (\alpha(1 - \cos\vartheta)]^2} \left(1 + \cos^2\vartheta\right) \times$$

$$\times \left(1 + \frac{\alpha^2(1 - \cos\vartheta)}{(1 + \cos\vartheta)[1 + \alpha(1 - \cos\vartheta)]}\right) \tag{13}$$

where $r_0$ is the classical electron radius, $\vartheta$ is the scattering angle and $\alpha$ is the fine-structure constant. The energy spectra of the emitted electron is a continuum (Compton continuum, see fig. 11) with a peak at its maximum value. This value is a fraction of the energy of the incoming photon and depends solely on it. The maximum can be straigthforwardly calculated from the above formula for $\vartheta = 0$. The angle of electron recoil is forward at $\vartheta = 0$ while the scattered photon will be aimed straight back. Their energy will be:

$$E_{e(\mathrm{max})} = h\nu - \frac{h\nu}{1 - 2\alpha} \tag{14}$$

$$h\nu'_{\mathrm{min}} = \frac{h\nu}{1 - 2\alpha} \tag{15}$$

For low-energy photons, when the scattering interaction takes place, little energy is transferred regardless of the probability of such an interaction. As the energy increases, the fractional transfer increases,

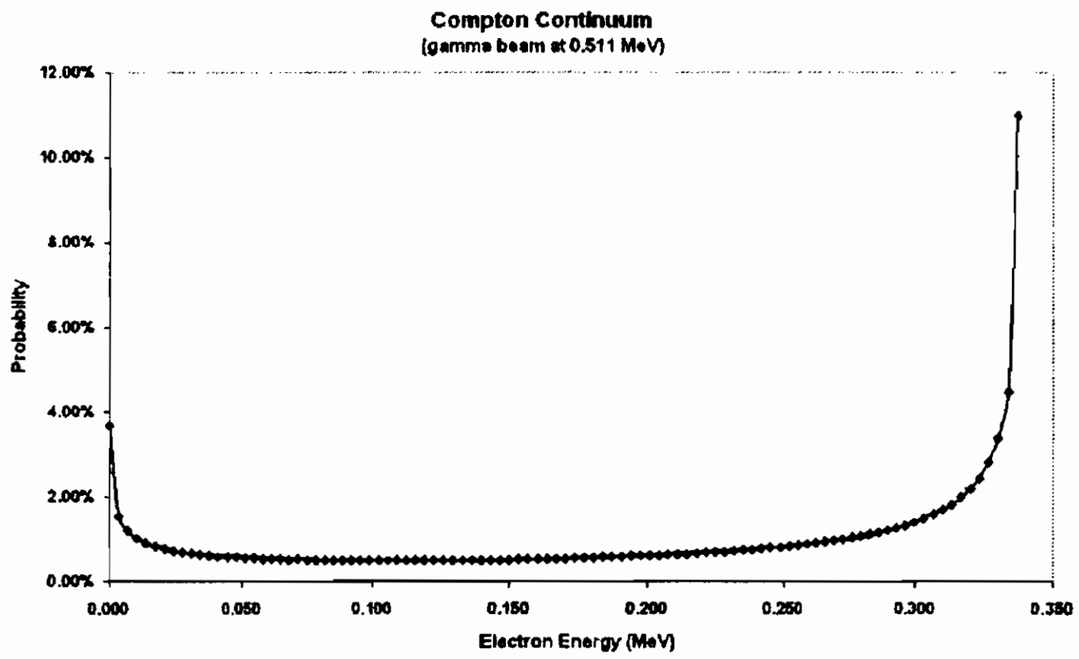



Figure 11: Compton continuum for a beam of 0.511 MeV $\gamma$.

approaching 1.0 for photons at energies above 10 to 20 MeV. For a monoenergetic gamma beam produced by electron-positron annihilation the maximum $E_e$ is about 340 KeV.

## 4.3 C: Electron interaction

The concept of range in matter is much less defined for electrons than for heavy charged particles because the electron total path length is much longer then the penetration distance along the initial velocity vector. For equivalent kinetic energy, the energy loss of electrons is much lower than that of heavy charged particles, so their path lenght is hundreds of times longer. To first approximation, the product of the range times the density of the absorber is a constant for different materials for electron of equal initial energy.

The energy collisional loss is given by:

$$-(\frac{dE}{dx})_c = \frac{2\pi e^4 NZ}{m_0 v^2}\left(\ln\frac{m_0 v^2 E}{2I^2(1-\beta^2)} - \ln 2(2\sqrt{1-\beta^2}-1-\beta^2)\right.$$
$$\left.+(1-\beta^2)\frac{1}{8}(1-\sqrt{1-\beta^2})^2\right) \qquad (16)$$

while the energy radiative loss is:

$$-(\frac{dE}{dX})_r = \frac{NEZ(Z+1)e^4}{137m_0^2c^4}\left(4\ln\frac{2E}{m_0c^2} - \frac{4}{3}\right) \qquad (17)$$

where $m_0$ is the electron mass, $N$ is the Avogadro number, $v$ is the electron speed and $\beta = v/c$.

# 5 Acknowledgments

The author is grateful to Ministero della Istruzione, Università e Ricerca (MIUR), the University of Pavia and Istituto Nazionale di Fisica Nucleare for cofinancing the PRIN project n. 2003-02-9728. He also acknowledges the members of the group: S.P. Ratti, P. Vitulo, M.M. Necchi, D. Pagano and C. Viviani for usefull and critical discussions.

# References

[1] E.A. Babichev et al.: Nucl. Instr. and Meth. **A419** (1998) 290.
[2] M. Lohmann et al.: Nucl. Instr. Meth. **A419** (1998) 276.
[3] a- A. Jeavons et al.: Nucl. Instr. Meth. **124** (1975) 491; b- A. Jeavons et al.: IEEE Trans. Nucl. Sci. **NS-46** (1999) 468.
[4] P. Bruyndonckx et al.: Nucl. Instr. Meth. **A392** (1997) 407.
[5] A. Brahme et al.: Nucl. Instr. Meth. **A454** (2000) 136. **NS-45** (1998) 1225.
[6] S. E. Derenzo: in *Positron Annihilation* Ed. The Japan Institute of Metals (1979).
[7] R. Santonico and R. Cardarelli: Nucl. Instr. Meth. **187** (1981) 377.
[8] F. W. K. Firk: Nucl. Instr. Meth. **A297** (1990) 532.
[9] F. H. Fahey: Radiol. Clin. N. Am. **39** (2001)919.
[10] W. W. Moses, S. E. Derenzo, P. Schlichta: Trans. Nucl. Sci. (1992) **NS-39** (1992) 1190.
[11] M.Abbrescia et al.: Nucl. Instr. Meth **A533** (2004) 7.

[12] D. F. Yu, J. A. Fessler: Nucl. Instr. Meth **A488** (2002) 362.
[13] J. A. Sorenson, M. E. Phelps: *Physics in nuclear medicine*, Ed. Sanders (1987).
[14] Report 52: *Particle counting in radioactivity measurements*, Ed. International Commission on Radiation Units and Measurements (1994).
[15] D.F. Yu, J.A. Fessler: Phys. Med. Biol. **45** (2000) 2043.
[16] J. W. Stayman, J. A. Fessler: IEEE Trans. Med. Imag. **23** (2004) 269.
[17] B. Karuta, R. Lecomte: IEEE Trans. Med. Imag. **11** (1992) 379.
[18] P. R. G. Virador, W. W. Moses, R. H. Huesman: IEEE Trans. Nucl. Sci. **NS-45** (1998) 1225.
[19] J. A. Fessler, W. L. Rogers: IEEE Tr. Im. Proc. **5** (1996) 1346.
[20] J. S. Liow, S. C. Strother: Phys. Med. Biol. **38** (1993) 55.
[21] R. Yao, J. Siedel, C. A. Johnson, M. E. Daube-Witherspoon, M. V. Green, R. E. Carson: IEEE Trans. Med. Imag. **19** (2000) 798.
[22] National Institute of Standards and Technology. Reference data: Athoms and molecular data; X-ray and $\gamma$-ray data; Radio dosimetry data (NIST website).
[23] D. L. Snyder, M. I. Miller, L. J. Thomas, D. G. Politte: IEEE Trans. Med. Imag. **6** (1987) 228.
[24] J. W. Stayman, J. A. Fessler: IEEE Trans. Med. Imag. **19** (2000) 601.
[25] J. W. Stayman, J. A. Fessler: in *Proceedings of the 1999 International Meeting on Fully 3D Image Reconstruction In Radiation Nuclear Medicine* (1999).
[26] S. E. Derenzo, R. H. Huesman, J. L. Cahoon, A. B. Geyer, W. W. Moses, D. C. Uber, T. Vuletich and T. F. Budinger: IEEE Trans. Nucl. Sci. **NS-35** (1988) 659.
[27] Z. H. Cho, J. Chan, L. Eriksson et al: J. Nucl. Med. **16** (1975) 1174.
[28] S. E. Derenzo: IEEE Trans. Nucl. Sci. **NS-33** (1986) 565.
[29] S. F. Haber, S. E. Derenzo, D. Uber: IEEE Trans Nucl Sci **NS-37** (1990) 1293.

[30] C.S. Levin, E. J. Hoffman: Phys. Med. Biol. **44** (1999) 781; also ref. 7.
[31] P. Colombino, B. Fiscella and L. Trossi: Nuovo Cimento **38** (1965) 707.
[32] G. Tarantola et al: J. Nucl. Med. **44** (2003) 756-769
[33] M. Yavuz, J. A. Fessler, in: Inform. Proces. Med. Imag. (Ed. Spinger, 1997) p. 190.
[34] M. E. Casey, E. J. Hoffman: J. Comp. Assisted Tomo. **10** (1986) 845.
[35] J. M. Ollinger, J. A. Fessler: IEEE Sig. Proc. Mag. **14** (1997) 43.
[36] E. J. Hoffman, S. C. Huang, M. E. Phelps, D. E. Kuhl: J. Comp. Assisted Tomo. **5** (1981) 391.
[37] E. U. Mumcuoglu, R. M. Leahy, S. R. Cherry: Phys. Med. Biol. **41** (1996) 1777.
[38] D. G. Politte, D. L. Snyder: IEEE Trans. Med. Imag. **10** (1991) 82.
[39] M. M. Ter-Pogossian, N. A. Mullani, D. C. Ficke et al: J. Comput. Assist. Tomo. **5** (1981) 227.
[40] M. M. Ter-Pogossian, D. C. Ficke, M. Yamamoto, J. T. Hood: in *Proc. Work. Time-of-Flight Pos. Tom.* (1982) p. 37.
[41] R. Gariod, R. Allemand, E. Cormoreche: *Proc. Work. Time-of-Flight Pos. Tom.* (1982).
[42] N. A. Mullani, J. Gaeta, K. Yerian et al: IEEE Trans. Nucl. Sci. **NS-31** (1984) 609.
[43] C. J. Thompson: IEEE Trans. Nucl. Sci. **NS-36** (1989) 1072.
[44] N. Mullani, J. Markham, M. Ter-Pogossian: J. Nucl. Med. **21** (1980) 1095.
[45] N. Mullani, W. Wong, R. Hartz et al: IEEE Trans. Nucl. Sci. **NS-29** (1982) 479.
[46] T. K. Lewellen, R. L. Harrison, A. N. Bice: IEEE Trans. Nucl. Sci. **NS-36** (1989) 1095.
[47] M. Yamamoto, N. Nohara, E. Tanaka, T. Tomitani et al: IEEE Trans. Nucl. Sci. **NS-36** (1989) 998.
[48] A. Mallon, P. Grangeat, P. X. Thomas: IEEE Nuclear Science Symposium Conference Record 92CH3232-6 **2** (1992) 988.