

# Kinetic Theory of Fluids

Henk van Beijeren  
Institute for Theoretical Physics  
Utrecht University

# Contents

<b>1</b>	<b>The Boltzmann Equation</b>	<b>3</b>
1.1	Derivation . . . . .	3
1.2	The Loss Term: $\Gamma_-$ . . . . .	4
1.2.1	The <i>Stosszahlansatz</i> . . . . .	5
1.3	The Gain Term: $\Gamma_+$ . . . . .	7
1.4	Interactions with the Walls: $\Gamma_W$ . . . . .	15
1.5	The Boltzmann Transport Equation . . . . .	19
<b>2</b>	<b>The <math>H</math>-Theorem of Boltzmann</b>	<b>19</b>
2.1	Some Models for the Interactions of the Gas Particles with the Boundaries . . . . .	21
2.2	Proof of the $H$ -Theorem . . . . .	22
2.3	The Equilibrium Distribution Function . . . . .	28
2.4	The Objections of Zermelo and of Loschmidt . . . . .	30
2.5	The Kac Ring Model. What is the Source of Irreversibility in the Boltzmann Equation? . . . . .	33
2.6	The Chapman-Enskog Solution . . . . .	37
2.6.1	The $\mu$ -expansion of the Boltzmann equation . . . . .	40
2.6.2	The distribution function to zeroth order in $\mu$ and the Euler equations . . . . .	41
2.6.3	The distribution function to order $\mu$ : The Navier-Stokes equations . . . . .	42
<b>3</b>	<b>The Liouville and BBGKY Hierarchy Equations</b>	<b>44</b>
3.1	Low densities. The Boltzmann closure . . . . .	46
3.2	Hard spheres. Enskog and revised Enskog closure . . . . .	48
<b>4</b>	<b>Enskog and Revised Enskog Equation</b>	<b>51</b>
4.1	Maximum entropy ensembles . . . . .	51
4.2	$H$ -theorem for the Revised Enskog Equation . . . . .	53
<b>5</b>	<b>Mode coupling and long time behavior of hydrodynamics</b>	<b>57</b>
5.1	Fluctuating Burgers equation[21] . . . . .	57
5.2	The Kardar-Parisi-Zhang equations[27] . . . . .	62
5.3	Scaling behavior . . . . .	63
5.4	The polynuclear growth model. Exact results . . . . .	64

# 1 The Boltzmann Equation

## 1.1 Derivation

The derivation of the Boltzmann equation was one of the great accomplishments of 19th century physics. Up to now it is still one of the very few fairly complete theories describing the time evolution out of equilibrium of a class of realistic and nontrivial physical systems, namely dilute gases and gas mixtures. Since the Boltzmann equation sets the scene for all kinetic equations describing many-particle systems I will start by reviewing what I consider as its main aspects.

The derivation of the Boltzmann equation proceeds by utilizing the fact that free streaming and collisions are the mechanisms by means of which the number of particles in a small volume  $\delta\mathbf{r}_1\delta\mathbf{v}_1$  about  $\mathbf{r}_1, \mathbf{v}_1$  is increased or decreased<sup>1</sup>. That is, the number density of such particles obeys an equation of the form

$$\begin{aligned} \frac{\partial f(\mathbf{r}_1, \mathbf{v}_1, t)}{\partial t} &= -\nabla_{\mathbf{r}_1} \cdot (\mathbf{v}_1 f(\mathbf{r}_1, \mathbf{v}_1, t)) - \nabla_{\mathbf{v}_1} \cdot (\dot{\mathbf{v}}_1 f(\mathbf{r}_1, \mathbf{v}_1, t)) + \\ &+ \Gamma_+ - \Gamma_- + \Gamma_W. \end{aligned} \quad (1)$$

The first two terms on the right hand side of Eq. (1) represent the change in  $f$  due to particles moving into and out of the small region without undergoing

---

<sup>1</sup>Regarding this region one may take two points of view: the traditional approach is assuming this region is not really infinitesimal, but its (spatial) diameter is small on the length scale of the mean free path between collisions, but large compared to the mean distance between particles, so it typically contains many particles. Equally, the time interval  $\delta t$  after which changes are monitored should be small compared to the mean free time between collisions of the same particle, but still so large that within the volume  $\delta\mathbf{r}_1\delta\mathbf{v}_1$  many particles stream in and out and several collisions occur during  $\delta t$ . The second interpretation is that the volume and the time change are infinitesimal and the numbers of particles (which then also are infinitesimal) are averages over a suitable non-equilibrium ensemble, for which  $f(\mathbf{r}_1, \mathbf{v}_1, t)$  describes the average time behavior of the particle distribution. The first interpretation becomes problematic for large values of the velocity, where the density of particles becomes very small, but has the advantage that it refers directly to individual systems. To apply the second interpretation to individual systems one has to argue that fluctuations of individual systems about the average behavior are typically small. One may actually account for these by adding fluctuating terms to the Boltzmann equation.

collisions and responding only to whatever external forces are present due to gravitational, electromagnetic, or other external fields (except for wall potentials that are accounted for by  $\Gamma_W$ ). The terms denoted by  $\Gamma_+$ ,  $\Gamma_-$  and  $\Gamma_W$  denote the rates of change of  $f$  due to respectively restituting collisions, loss collisions and collisions of gas particles with the walls of the container. If these three collision terms were neglected, the resulting equation would simply be an example of the differential expression for the conservation of the total number of particles in the  $2d$ -dimensional position and velocity space. However, the change in the distribution function due to collisions is the essential part of the Boltzmann equation, and we now turn our attention to obtaining expressions for  $\Gamma_+$ ,  $\Gamma_-$  and  $\Gamma_W$ .

## 1.2 The Loss Term: $\Gamma_-$

The expression for  $\Gamma_-$  requires less analysis of the collision dynamics than  $\Gamma_+$  but it does require the same essential assumptions as do the other terms. We need to determine the rate at which particles with velocity  $\mathbf{v}_1$  are lost in a region  $\delta\mathbf{r}_1\delta\mathbf{v}_1$  about the point  $\mathbf{r}_1, \mathbf{v}_1$  at time  $t$  through binary encounters with other gas particles. We will calculate  $\Gamma_-$  for a system of particles in three spatial dimensions, and the extension to other dimensions will be immediate.

Let us consider the collisions that take place between particles with velocity  $\mathbf{v}_1$ , called particle 1, and those with some other velocity, say  $\mathbf{v}_2$ , called particle 2. We refer to these collisions as  $(\mathbf{v}_1, \mathbf{v}_2)$  collisions. We now pick one of the particles with velocity  $\mathbf{v}_1$  and imagine that it is at the center of a coordinate system with  $z$ -axis in the direction of  $\mathbf{v}_2 - \mathbf{v}_1$  as illustrated in Fig. 1. A collision takes place whenever the center of the particle with velocity  $\mathbf{v}_2$  is within a distance  $a$  from the origin of the coordinate system. For central forces the relative motion takes place in a plane that includes the  $z$ -axis in the direction of  $\mathbf{v}_2 - \mathbf{v}_1$  and the trajectory of particle 2 before its center reaches the sphere of radius  $a$ , called the *action sphere* about particle 1. We measure the rotation of the collision plane about a fixed plane that includes the  $z$ -axis by an azimuthal angle  $\epsilon$ . The final quantity that characterizes the collision, given  $\mathbf{v}_1, \mathbf{v}_2$  and  $\epsilon$  is the impact parameter  $b$ , which is the distance between the trajectory of particle 2 before the collision and the  $z$ -axis. For  $b > a$  no collision can take place since the center of particle 2 never enters the action sphere. Suppose now that the impact parameter  $b$  is such that  $b < a$ . By looking at Fig. 1 we can see that if the center of particle 2 is inside the *collision cylinder* with volume  $\pi a^2|\mathbf{v}_2 - \mathbf{v}_1|\delta t$ , a  $(\mathbf{v}_1, \mathbf{v}_2)$  collision

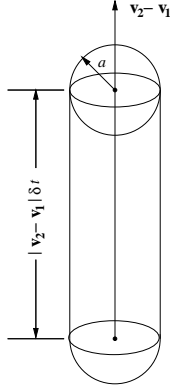


Figure 1: The  $(\mathbf{v}_1, \mathbf{v}_2)$ -collision cylinder. The sphere has a radius  $a$  which is the range of the forces. For hard sphere molecules,  $a$  is the diameter of the molecules.

will take place within time interval  $\delta t$ , if no other particles intervene earlier, which we assume to be the case. We also ignore any collisions that may be already in progress at time  $t$ . Such events lead to higher order density contributions, and vanish if the particles are hard spheres with instantaneous binary collisions.

### 1.2.1 The *Stosszahlansatz*

We now have all the geometry that we need to begin assembling the components needed to obtain Boltzmann's expression for  $\Gamma_-$ . The argument proceeds as follows:

- The number of particles with velocity  $\mathbf{v}_1$  in the volume  $\delta\mathbf{r}_1\delta\mathbf{v}_1$  at time  $t$  is  $f(\mathbf{r}_1, \mathbf{v}_1, t)\delta\mathbf{r}_1\delta\mathbf{v}_1$ .
- Each collision cylinder attached to a particle with velocity  $\mathbf{v}_1$  for a collision with a particle with velocity  $\mathbf{v}_2$  has volume  $\pi a^2|\mathbf{v}_2 - \mathbf{v}_1|\delta t$ , so that the total volume of these collision cylinders in  $\delta\mathbf{r}_1\delta\mathbf{v}_1$  is  $f(\mathbf{r}_1, \mathbf{v}_1, t)\delta\mathbf{r}_1\delta\mathbf{v}_1 \times$

$$\pi a^2|\mathbf{v}_2 - \mathbf{v}_1|\delta t.$$

- In principle, we now would have to examine each of these collision cylinders to see if one of them is occupied by a particle with velocity

$\mathbf{v}_2$ . Since we are unable to do this without following the dynamical history of each and every particle in the system<sup>2</sup>, we need to use some probabilistic reasoning to calculate the (average) number of  $(\mathbf{v}_1, \mathbf{v}_2)$  collisions.

- The stochastic argument is based upon Boltzmann's *Stosszahlansatz*: The number of  $(\mathbf{v}_1, \mathbf{v}_2)$  collisions in the volume  $\delta\mathbf{r}_1\delta\mathbf{v}_1\delta\mathbf{v}_2$  in time  $\delta t$  is given by the product of the total volume of the collision cylinders under consideration and the average number of particles per unit volume with velocity  $\mathbf{v}_2$  in the range  $\delta\mathbf{v}_2$ , namely  $f(\mathbf{r}_1, \mathbf{v}_2, t)\delta\mathbf{v}_2$ . This product is

$$f(\mathbf{r}_1, \mathbf{v}_1, t)f(\mathbf{r}_1, \mathbf{v}_2, t)\delta\mathbf{r}_1\delta\mathbf{v}_1\delta\mathbf{v}_2\pi a^2|\mathbf{v}_2 - \mathbf{v}_1|\delta t. \quad (2)$$

Now we have assembled all of the pieces needed for the calculation of  $\Gamma_-$ . It is obtained by integrating the number of  $(\mathbf{v}_1, \mathbf{v}_2)$  collisions just obtained over all possible velocities  $\mathbf{v}_2$ . That is

$$\Gamma_- \delta\mathbf{r}_1\delta\mathbf{v}_1\delta t = \pi a^2 \int d\mathbf{v}_2 |\mathbf{v}_1 - \mathbf{v}_2| f(\mathbf{r}_1, \mathbf{v}_1, t)f(\mathbf{r}_1, \mathbf{v}_2, t)\delta\mathbf{r}_1\delta\mathbf{v}_1\delta t. \quad (3)$$

The expression for  $\Gamma_-$  is then given by

$$\Gamma_- = \pi a^2 \int d\mathbf{v}_2 |\mathbf{v}_1 - \mathbf{v}_2| f(\mathbf{r}_1, \mathbf{v}_1, t)f(\mathbf{r}_1, \mathbf{v}_2, t). \quad (4)$$

Some additional assumptions were made in the final steps. First by setting the number of  $\mathbf{v}_1, \mathbf{v}_2$  collisions equal to the number of particles with velocity in the volume  $\delta\mathbf{v}_2$  around  $\mathbf{v}_2$  in the  $(\mathbf{v}_1, \mathbf{v}_2)$  collision cylinder, we assumed that any particle with velocity  $\mathbf{v}_2$  in such a cylinder indeed collides with the particle with velocity  $\mathbf{v}_1$  to which the cylinder is attached. This can certainly be defended in the limit  $\delta t \rightarrow 0$ . It was also assumed that the distribution functions do not vary significantly between  $\mathbf{r}_1$  and the actual position  $\mathbf{r}_2$  of the other colliding particle, so that both distribution functions can be evaluated at the same spatial point,  $\mathbf{r}_1$ .<sup>3</sup>

We emphasize that the *Stosszahlansatz* is a probabilistic statement, not a mechanical one. It is a statement about the chance of finding a particle

---

<sup>2</sup>This process is indeed carried out in computer simulated molecular dynamics.

<sup>3</sup>We will see how to account for changes in the distribution function over distances of order  $a$  in the derivation of the Enskog equation.

with velocity  $\mathbf{v}_2$  in a collision cylinder. That is, the *Stosszahlansatz* is the assumption that, before a collision, particles about to collide are completely uncorrelated, so that, irrespective of the knowledge that there is a particle with velocity  $\mathbf{v}_1$  at position  $\mathbf{r}_1$ , the number of particles with velocity in the volume  $\delta\mathbf{v}_2$  around  $\mathbf{v}_2$  contained in the collision cylinders is taken to be equal to product of *the number of particles with velocity in the volume  $\delta\mathbf{v}_2$  around  $\mathbf{v}_2$  per unit volume* with *the volume of the collision cylinder*. In realistic systems correlations between colliding particles are indeed important since, among other things, they are responsible for mode-coupling effects giving rise to slow decay of current-current time correlation functions, and these correlations must be included in any systematic extension of the Boltzmann equation to gases at higher densities.

### 1.3 The Gain Term: $\Gamma_+$

Next we turn our attention to the calculation of  $\Gamma_+$ , the rate at which particles with velocity  $\mathbf{v}_1$  are created in the small volume  $\delta\mathbf{r}_1\delta\mathbf{v}_1$  about  $(\mathbf{r}_1, \mathbf{v}_1)$ . The task at hand is to find the velocities of two colliding particles such that after collision one of them has velocity  $\mathbf{v}_1$ . We will then use arguments similar to those used to obtain  $\Gamma_-$  to obtain an expression for  $\Gamma_+$ . However, the calculation of  $\Gamma_+$  requires a more careful discussion of the collision dynamics since this term depends on a determination of *restituting* velocities. These are the velocities that a pair of colliding particles must have before a collision so that one of them will have velocity  $\mathbf{v}_1$  after the collision.

First we examine the *direct* collision, that is, the  $\mathbf{v}_1, \mathbf{v}_2$  collision discussed above. Since we assume that the particles interact with central, conservative forces, there are three conservation laws that must be satisfied, namely, the conservation of momentum, that of energy, and that of angular momentum. The conservation of energy and of momentum may be stated in terms of the initial velocities  $\mathbf{v}_1, \mathbf{v}_2$ , and the velocities after the collision,  $\mathbf{v}'_1, \mathbf{v}'_2$  as

$$\mathbf{v}_1 + \mathbf{v}_2 = \mathbf{v}'_1 + \mathbf{v}'_2 \tag{5}$$

$$v_1^2 + v_2^2 = v_1'^2 + v_2'^2, \tag{6}$$

where we used the fact that the two colliding particles are mechanically identical to eliminate the mass in these conservation equations. An immediate consequence of these equations is the conservation of the magnitude of the

relative velocity,  $\mathbf{g} = \mathbf{v}_2 - \mathbf{v}_1$ ,

$$g = |\mathbf{v}_2 - \mathbf{v}_1| = |\mathbf{v}'_2 - \mathbf{v}'_1| = g'. \quad (7)$$

This follows from the fact that the energies before and after collision are purely kinetic energies and may be expressed as the sum of the kinetic energy of the center of mass and the relative kinetic energy, as

$$v_1^2 + v_2^2 = 2mV^2 + \frac{1}{2}g^2 \quad (8)$$

where  $\mathbf{V} = (\mathbf{v}_1 + \mathbf{v}_2)/2$  is the velocity of the center of mass. Since, the kinetic energy of the center of mass is conserved separately as a result of the conservation of momentum, the relative kinetic energy is also conserved, from which (7) follows immediately.

Conservation of angular momentum will have a somewhat more delicate role to play in the determination of the restituting velocities as we will see below. First we set up the geometry of the direct collision in the relative coordinate system with origin fixed at the center of particle 1. We consider the relative motion, which takes place in the plane through the origin normal to the relative angular momentum  $m\mathbf{r}_{21} \times \mathbf{g}/2$ . Before the collision, the center of particle 2 moves toward particle 1 in a straight line with velocity  $\mathbf{g}$ . When it reaches the action sphere of radius  $a$  about the origin, the collision begins. After the collision is completed, particle 2 will again move along a straight line, but with velocity  $\mathbf{g}'$ . Let us denote the distance of this line from a parallel line through the origin by  $b'$ . The conservation of angular momentum requires that  $gb = g'b'$ , but the fact that  $g = g'$  implies that  $b' = b$ , as illustrated in Fig. 2.

If we follow the collision through it's completion, we can see that the spatial trajectory of particle 2, has a mirror symmetry about a line, called the *apse line*, from the origin through the point of closest approach of the trajectory to the origin. There are a number of ways to prove this symmetry, but the simplest is to imagine two particles, the one we have been discussing and a "ghost particle" that travels toward the origin on the other side of the apse line, along the trajectory, but in the opposite direction. If the two particles start at equal distances from the action sphere, with the same impact parameters, and the same speed, they will each experience central forces of equal magnitude at equal distances from the origin and will arrive at the apse line at the same time with exactly opposite velocities. At that

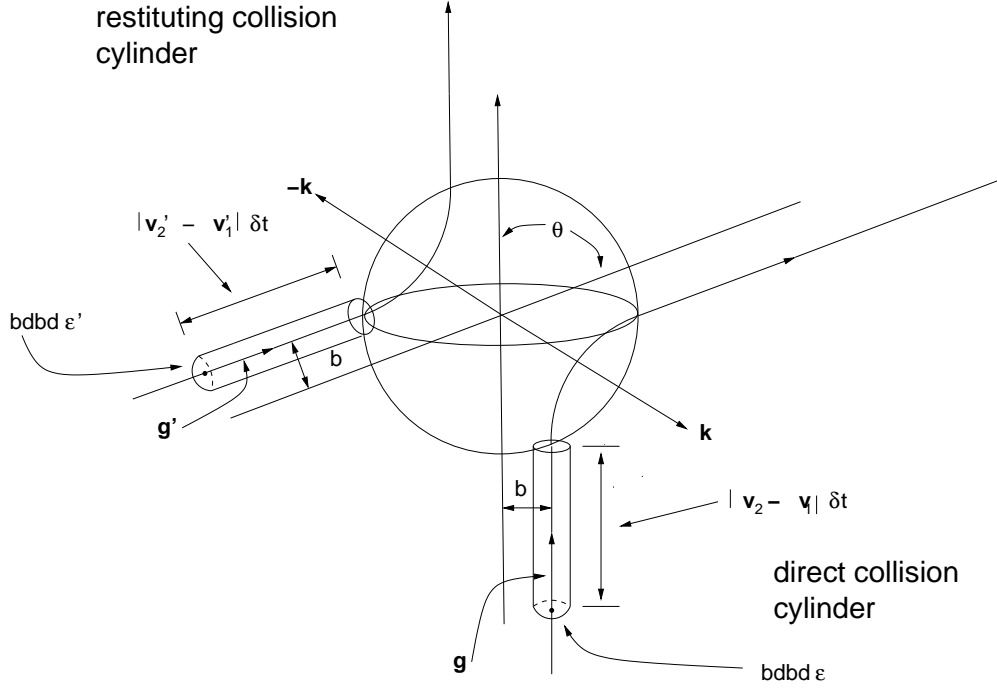


Figure 2: Direct and restituting collisions in the relative coordinate frame. The corresponding collision cylinders, as well as the scattering angle,  $\theta$ , are illustrated.

point the two trajectories will be mirror images of each other through the apse line. Based on this observation, we see that the velocities  $\mathbf{g}$  and  $\mathbf{g}'$ , have equal components in the direction perpendicular to the apse line, and opposite components in the direction of the apse line. If we denote by  $\hat{\sigma}$  a unit vector in the direction to the point of closest approach from the origin along the apse line, it follows from this discussion that

$$\mathbf{g}' = \mathbf{g} - 2(\mathbf{g} \cdot \hat{\sigma}) \hat{\sigma}. \quad (9)$$

This result allows us to express the final velocities of both particles in terms of the initial relative velocity and the unit vector  $\hat{\sigma}$  along the apse line, as

$$\begin{aligned} \mathbf{v}'_1 &= \mathbf{v}_1 + (\mathbf{g} \cdot \hat{\sigma}) \hat{\sigma}, \text{ and} \\ \mathbf{v}'_2 &= \mathbf{v}_2 - (\mathbf{g} \cdot \hat{\sigma}) \hat{\sigma}. \end{aligned} \quad (10)$$

Of course, the crucial quantity that determines the outcome of the collision is the unit vector  $\hat{\sigma}$ , which, in turn, depends upon the impact parameter  $b$ ,

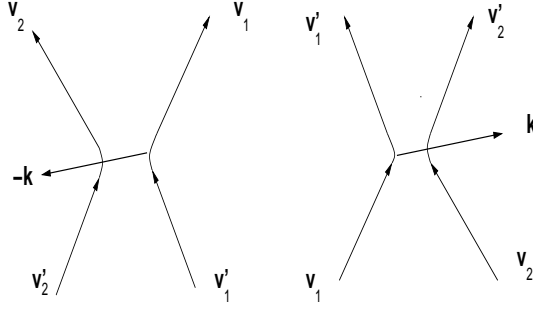


Figure 3: Schematic illustration of the direct collisions, on the right, and the restituting collisions, on the left. The corresponding unit vectors indicating the direction of the apse lines are also shown.

and upon both the direction and the magnitude of  $\mathbf{g}$ . The actual form of this relationship is determined by the interaction potential between the two particles. It is at this point, the determination of the direction of  $\hat{\sigma}$ , that one needs more information about the collision dynamics than is provided by the conservation laws.

Assuming we can solve the problem of determining  $\hat{\sigma}$  as function of  $b$  and  $\mathbf{g}$  we are now in a position to determine the restituting velocities. We imagine two particles that are aimed to collide with the same impact parameter  $b$  as in the direct collision, and with pre-collision velocities given by  $\mathbf{v}'_1$  and  $\mathbf{v}'_2$  (but with opposite relative angular momentum). This collision is illustrated in Fig. 3 and it can easily be visualized by a simple rearrangement of the particles in space. Under these circumstances the apse line of this new collision will be in the direction of  $-\hat{\sigma}$ , opposite to that of the direct collision. Notice now that the relative velocity after this collision will be the original relative velocity, and we have succeeded in determining the restituting velocities, in terms of the parameters and results of the direct collision. To see this we denote, for the moment, the relative velocity after this second type of collision by  $\mathbf{g}''$  and use Eq. (9) to obtain

$$\mathbf{g}'' = \mathbf{g}' - 2(\mathbf{g}' \cdot \hat{\sigma}) \hat{\sigma} = \mathbf{g} - 2(\mathbf{g} \cdot \hat{\sigma}) \hat{\sigma} + 2(\mathbf{g} \cdot \hat{\sigma}) \hat{\sigma} = \mathbf{g}. \quad (11)$$

It follows from Eq. (11) that one can determine the restituting velocities simply by finding the final velocities for the direct collisions. These velocities depend, as we have seen, on the collision parameters and on the interaction potential. The observation that the final velocities of the direct collisions are

the initial velocities of the restituting collisions and *vice versa* will play an essential role when we turn our attention to the Boltzmann H-Theorem later in this section.

We now have enough information to determine  $\Gamma_+$ . We use arguments similar to those used to obtain  $\Gamma_-$ , but with a more delicate construction of the collision cylinders. Since the direction of the unit vector,  $\hat{\boldsymbol{\sigma}}$ , along the apse line depends upon the parameters of the collision, we consider collision cylinders for all possible values of the collision parameters. We illustrate the calculation for three dimensional systems and the extensions to other spatial dimensions will be immediate. For three dimensions the collision parameters of the direct collisions are the impact parameter,  $b$ , and the azimuthal angle  $\epsilon$  of the collision plane. The latter is a polar angle in a plane perpendicular to  $\mathbf{g}$  measured with respect to a fixed plane in the relative coordinate system that contains the  $z$ -axis, taken to be in the direction of  $\mathbf{g}$ . For the restituting collisions, the collision parameters are the impact parameters, taken, for reasons given above, to be the same as in the direct collision for each pair of final velocities  $\mathbf{v}'_1, \mathbf{v}'_2$ , but the azimuthal angle,  $\epsilon'$  is a polar angle in a plane perpendicular to  $\mathbf{g}'$  measured with respect to a fixed plane containing the  $z'$ -axis which is along the direction of  $\mathbf{g}'$ .

Let us now consider restituting collisions that take place with impact parameters in the range  $b$  to  $b+db$ , and azimuthal angles in the range between  $\epsilon'$  and  $\epsilon' + d\epsilon'$ . We follow the method used above to obtain  $\Gamma_-$ , *viz.*

1. The volume of a small collision cylinder about particle 1 for  $\mathbf{v}'_1, \mathbf{v}'_2$  collisions is  $b|\mathbf{g}'|\delta t db d\epsilon'$ ;
2. The number of such collision cylinders in the 6-dimensional volume  $\delta\mathbf{r}_1\delta\mathbf{v}'_1$  is  $f(\mathbf{r}_1, \mathbf{v}'_1, t)\delta\mathbf{r}_1\delta\mathbf{v}'_1$  which is simply the number of particles with velocity  $\mathbf{v}'_1$  in this volume;
3. We now apply the *Stosszahlansatz* for the restituting collisions: The number of particles with velocity in the volume  $\delta\mathbf{v}'_2$  around  $\mathbf{v}'_2$  in these collision cylinders is the product of the total volume of all the  $\mathbf{v}'_1, \mathbf{v}'_2$  collision cylinders considered above with the number of particles per unit volume with velocities in the range  $\delta\mathbf{v}'_2$  about velocity  $\mathbf{v}'_2$ , at time  $t$ , given by  $f(\mathbf{r}_1, \mathbf{v}'_2, t)\delta\mathbf{v}'_2$ .
4. We assume that all of the particles with velocity  $\mathbf{v}'_2$  in the collision cylinders actually collide with the particles with velocity  $\mathbf{v}'_1$ .

Putting all of these pieces together we find that the number of particles produced by  $\mathbf{v}'_1\mathbf{v}'_2$  collisions with velocity  $\mathbf{v}_1$  in the region  $\delta\mathbf{r}_1\delta\mathbf{v}'_1$  in time  $\delta t$  is given by

$$f(\mathbf{r}_1, \mathbf{v}'_1, t)f(\mathbf{r}_1, \mathbf{v}'_2, t)|\mathbf{g}'|bdbd\epsilon'\delta\mathbf{r}_1\delta\mathbf{v}'_1\delta\mathbf{v}'_2\delta t. \quad (12)$$

In order to put this expression into a form that can be used to obtain  $\Gamma_+$  we make use of some transformations of the volume elements in velocity space and of the collision parameters.

The expression for  $\Gamma_-$  as well as expression (12) can be written in terms of  $\hat{\boldsymbol{\sigma}}$ , the unit vector in the direction of the apse line, by using the fact that, in the collision plane,  $\hat{\boldsymbol{\sigma}}$  makes an angle of  $(\pi + \theta)/2$  with the  $z$ -axis, and  $-\hat{\boldsymbol{\sigma}}$  makes the same angle with respect to the  $z'$ -axis, as illustrated in Fig. 2. For the direct collisions we may write

$$\begin{aligned} d\hat{\boldsymbol{\sigma}} &= \sin\left[\frac{1}{2}(\pi + \theta)\right] d\left[\frac{1}{2}(\pi + \theta)\right] d\epsilon \\ &= \frac{1}{2}\cos(\theta/2)d\theta d\epsilon \\ &= \frac{1}{2}\cos(\theta/2)\left|\frac{d\theta}{db}\right|dbd\epsilon. \end{aligned} \quad (13)$$

Therefore we may write

$$|\mathbf{v}_2 - \mathbf{v}_1|bdbd\epsilon = B(\mathbf{g}, \hat{\boldsymbol{\sigma}})d\hat{\boldsymbol{\sigma}}, \quad (14)$$

where

$$B(\mathbf{g}, \hat{\boldsymbol{\sigma}}) = 2g\left|\frac{db}{d\theta}\right|[\cos(\theta/2)]^{-1}. \quad (15)$$

The quantity  $B(\mathbf{g}, \hat{\boldsymbol{\sigma}})$  depends upon the intermolecular potential since that determines the scattering angle,  $\theta$  for a given value of  $b$  and  $g$ , but  $B$  does not depend upon the azimuthal angle  $\epsilon$ . In a similar way, for the restituting collisions we can write

$$B(\mathbf{g}', -\hat{\boldsymbol{\sigma}})d\hat{\boldsymbol{\sigma}} = gbdbd\epsilon', \quad (16)$$

where we have used the fact that the apse line points in the direction of  $-\hat{\boldsymbol{\sigma}}$  for the restituting collisions. It is important to note that we have used  $d(-\hat{\boldsymbol{\sigma}}) = d\hat{\boldsymbol{\sigma}}$ , and the fact that the same unit vector  $\hat{\boldsymbol{\sigma}}$  can be described in a coordinate system with  $z$ -axis in the direction of  $\mathbf{g}$  or in a coordinate with

a  $z$ -axis in the direction of  $\mathbf{g}'$ , as in the restituting collisions. Furthermore, since  $|\mathbf{g}| = |\mathbf{g}'|$ , and  $-\hat{\boldsymbol{\sigma}}$  makes an angle of  $(\pi + \theta)/2$  with  $\mathbf{g}'$ , it follows that

$$B(\mathbf{g}', -\hat{\boldsymbol{\sigma}}) = B(\mathbf{g}, \hat{\boldsymbol{\sigma}}). \quad (17)$$

Equation (17) must be regarded with some caution. It is correct as it stands if  $b$  is a single valued function of  $\theta$ . As we will discuss later in this section (See Section 2.7), typical potentials that are repulsive at small inter-particle separations and attractive at larger separations, may have more than one value of the impact parameter  $b$  for a given scattering angle  $\theta$ . In that case one must re-define  $B(\mathbf{g}, \hat{\boldsymbol{\sigma}})d\hat{\boldsymbol{\sigma}}$  to be the sum of all possible values of  $gbdbd\epsilon$  for which the impact parameters  $b$  lead to the same  $\hat{\boldsymbol{\sigma}}$ , and a similar sum for  $B(\mathbf{g}', -\hat{\boldsymbol{\sigma}})d\hat{\boldsymbol{\sigma}}$ . Then with these definitions the equality Eq. (17) is correct.

Next we note that  $\Gamma_+ \delta \mathbf{r}_1 \delta \mathbf{v}_1 \delta t$  is given by

$$\Gamma_+ \delta \mathbf{r}_1 \delta \mathbf{v}_1 \delta t = \delta \mathbf{r}_1 \delta t \int d\hat{\boldsymbol{\sigma}} \int_{R'} d\mathbf{v}'_1 \int d\mathbf{v}'_2 B(\mathbf{g}, \hat{\boldsymbol{\sigma}}) f(\mathbf{r}_1, \mathbf{v}'_1, t) f(\mathbf{r}_1, \mathbf{v}'_2, t), \quad (18)$$

where  $R'$  denotes the region in the space of the variables  $\mathbf{v}'_1, \mathbf{v}'_2, \hat{\boldsymbol{\sigma}}$  determined by the requirement that particle 1 has velocity in the region  $\delta \mathbf{v}_1$  about velocity  $\mathbf{v}_1$  after the  $(\mathbf{v}'_1, \mathbf{v}'_2)$  collision. This integration is most easily performed if we note that the Jacobian for the transformation from restituting collision velocities to direct collision velocities is unity,

$$d\mathbf{v}'_1 d\mathbf{v}'_2 = d\mathbf{v}_1 d\mathbf{v}_2, \quad \text{for fixed } \hat{\boldsymbol{\sigma}}. \quad (19)$$

This is most easily seen by imagining a transformation from laboratory velocities to center of mass and relative velocities for the velocities before and after a collision. The center of mass velocity is unaffected by the collision, and the relative velocity is just reflected with respect to the plane normal to  $\hat{\boldsymbol{\sigma}}$ , as described by (9). Since the Jacobian of such a reflection is unity, Eq. (19) follows immediately. With this result we can write Eq. (18) as

$$\Gamma_+ \delta \mathbf{r}_1 \delta \mathbf{v}_1 \delta t = \delta \mathbf{r}_1 \delta t \int d\hat{\boldsymbol{\sigma}} \int_R d\mathbf{v}_1 \int d\mathbf{v}_2 B(\mathbf{g}, \hat{\boldsymbol{\sigma}}) f(\mathbf{r}_1, \mathbf{v}'_1, t) f(\mathbf{r}_1, \mathbf{v}'_2, t), \quad (20)$$

Where  $R$  denotes the region of  $\mathbf{v}_1, \mathbf{v}_2, \hat{\boldsymbol{\sigma}}$  space where  $\mathbf{v}_1, \mathbf{v}_2$  are the final velocities after the  $(\mathbf{v}'_1, \mathbf{v}'_2)$  collision with apse line in the direction of  $-\hat{\boldsymbol{\sigma}}$ , under the condition that the velocity of particle 1 is allowed to range over the small region,  $\delta \mathbf{v}_1$ . Note that we have set the final integration of the right

hand side of Eq. (20) to be the  $\hat{\sigma}$  integration since Eq. (19) holds only for fixed  $\hat{\sigma}$ . However we can now rearrange the integrals on the right hand side of Eq. (20) to obtain

$$\begin{aligned}\Gamma_+ \delta \mathbf{r}_1 \delta \mathbf{v}_1 \delta t &= \delta \mathbf{r}_1 \delta t \int_{\delta \mathbf{v}_1} d\mathbf{v}_1 \int d\mathbf{v}_2 \int d\hat{\sigma} B(\mathbf{g}, \hat{\sigma}) f(\mathbf{r}_1, \mathbf{v}'_1, t) f(\mathbf{r}_1, \mathbf{v}'_2, t), \\ &= \delta \mathbf{r}_1 \delta \mathbf{v}_1 \delta t \int d\mathbf{v}_2 \int d\hat{\sigma} B(\mathbf{g}, \hat{\sigma}) f(\mathbf{r}_1, \mathbf{v}'_1, t) f(\mathbf{r}_1, \mathbf{v}'_2, t),\end{aligned}\quad (21)$$

for sufficiently small  $\delta \mathbf{v}_1$ . The expression for  $\Gamma_+$  follows immediately. We have now obtained Boltzmann's collision integral, which by referring to Eqs. (1), (4) and (21) we can write as

$$\begin{aligned}\Gamma_+ - \Gamma_- &= \int d\mathbf{v}_2 \int d\hat{\sigma} B(\mathbf{g}, \hat{\sigma}) [f(\mathbf{r}_1, \mathbf{v}'_1, t) f(\mathbf{r}_1, \mathbf{v}'_2, t) \\ &\quad - f(\mathbf{r}_1, \mathbf{v}_1, t) f(\mathbf{r}_1, \mathbf{v}_2, t)].\end{aligned}\quad (22)$$

It is often convenient to use a shorthand notation whereby

$$\begin{aligned}f_1 &\equiv f(\mathbf{r}_1, \mathbf{v}_1, t), & f_2 &\equiv f(\mathbf{r}_1, \mathbf{v}_2, t); \\ f'_1 &\equiv f(\mathbf{r}_1, \mathbf{v}'_1, t), & f'_2 &\equiv f(\mathbf{r}_1, \mathbf{v}'_2, t),\end{aligned}\quad (23)$$

which enables us to write

$$\Gamma_+ - \Gamma_- = \int d\mathbf{v}_2 \int d\hat{\sigma} B(\mathbf{g}, \hat{\sigma}) [f'_1 f'_2 - f_1 f_2].\quad (24)$$

In addition we use the notation  $J(f, f)$  for this term. That is

$$J(f, f) \equiv \int d\mathbf{v}_2 \int d\hat{\sigma} B(\mathbf{g}, \hat{\sigma}) [f'_1 f'_2 - f_1 f_2]\quad (25)$$

It is worth repeating at this point that our derivation of the gain and loss terms has made some essential assumptions, namely

1. The gas consists of monatomic particles that interact with central, short ranged forces.
2. The gas is sufficiently dilute that only binary collisions need be taken into account.

3. The number of direct and restituting collisions taking place in a short time interval  $\delta t$  can be calculated using the *Stosszahlansatz*.
4. The distribution functions vary slowly enough with position that the difference between the positions of the centers of the colliding particles immediately before collision can be neglected when the rates of direct and restituting collisions are calculated.
5. The distribution functions vary slowly enough with time that their changes over a time interval on the order of the duration of a binary collision also be neglected.

#### 1.4 Interactions with the Walls: $\Gamma_W$

The final term in the form of the Boltzmann equation used here is somewhat unusual. In the usual derivations of this equation, the boundaries are not included in the equation but are accounted for by specifying boundary conditions that the distribution functions must satisfy for positions on the boundaries of the vessel. Here we treat the boundaries as if they were just another kind of particle with which the gas particles interact. While this may seem artificial for a gas in a container, the same method can be used to describe the motion of a large particle in a gas, such as would occur in a description of Brownian motion or of Stokes or rarefied gas flow around a sphere or other macroscopic object. Thus our method here can be seen as setting the stage for later applications of the Boltzmann equation, and of kinetic theory, in general.

As in the previous discussion we focus on particles with some velocity,  $\mathbf{v}_1$ , and write

$$\Gamma_W = \Gamma_W^{(+)} - \Gamma_W^{(-)}, \quad (26)$$

where  $\Gamma_W^{(+)}$  is the rate at which particles in region  $\delta\mathbf{r}_1$  acquire velocity  $\mathbf{v}_1$  through collisions with the boundaries, and  $\Gamma_W^{(-)}$  is the rate at which particles in region  $\delta\mathbf{r}_1$  with velocity  $\mathbf{v}_1$  collide with the boundary and thereby change their velocity. We begin, as before, with the loss term,  $\Gamma_W^{(-)}$ . To compute this term we fix our attention on a small section of the surface located with respect to some fixed coordinate system by  $\boldsymbol{\rho}_S$ , as illustrated in Fig. 4.

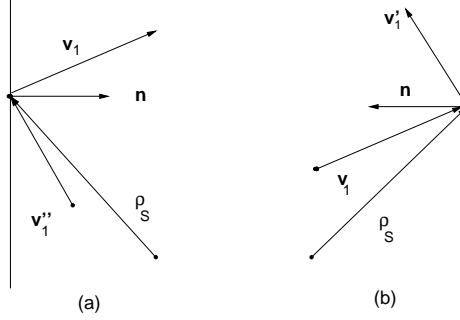


Figure 4: Schematic illustration of particle-wall collisions. In (a) the number of particles with velocity  $\mathbf{v}_1$  is increased due to collisions of particles with the wall. In (b) the number of particles with velocity  $\mathbf{v}_1$  is diminished when one of them collides with the wall.

Let  $dS$  be the area of the small section and let  $\hat{\mathbf{n}}$  be a unit vector normal to the surface and pointing into the region occupied by the gas. Only particles with velocities  $\mathbf{v}_1$  satisfying the condition  $\mathbf{v}_1 \cdot \hat{\mathbf{n}} < 0$  can collide with the wall. The probability that a particle at position  $\mathbf{r}_1$  in the range  $\delta\mathbf{r}_1$  is on the boundary at point  $\boldsymbol{\rho}_S$  and heading toward the wall is  $\delta(\boldsymbol{\rho}_S - \mathbf{r}_1)\delta\mathbf{r}_1\Theta(-\mathbf{v}_1 \cdot \hat{\mathbf{n}})$ . Here  $\delta(\mathbf{r}_1 - \boldsymbol{\rho}_S)$  is a Dirac delta function, and  $\Theta(x)$  is the Heaviside function taking the value unity when  $x > 0$  and zero otherwise. By constructing a small collision cylinder with base  $dS$  on the small surface element and height  $|\mathbf{v}_1 \cdot \hat{\mathbf{n}}|\delta t$ , we see that the number of particles that are removed from the velocity range  $\delta\mathbf{v}_1$  about velocity  $\mathbf{v}_1$ , in the volume element  $\delta\mathbf{r}_1$  about  $\mathbf{r}_1$  in time  $\delta t$  due to collisions with the boundary at  $\boldsymbol{\rho}_S$  is given by

$$\Gamma_W^{(-)}\delta t\delta\mathbf{v}_1\delta\mathbf{r}_1 = \int_{\partial V} dS|\mathbf{v}_1 \cdot \hat{\mathbf{n}}|\delta(\mathbf{r}_1 - \boldsymbol{\rho}_S)\Theta(-\mathbf{v}_1 \cdot \hat{\mathbf{n}})f(\mathbf{r}_1, \mathbf{v}_1, t)\delta t\delta\mathbf{r}_1\delta\mathbf{v}_1, \quad (27)$$

where we have carried out an integration over the entire surface bounding the gas, which we denote by  $\partial V$ . This expression leads directly to one for  $\Gamma_W^{(-)}$  as

$$\Gamma_W^{(-)} = \int_{\partial V} dS|\mathbf{v}_1 \cdot \hat{\mathbf{n}}|\delta(\mathbf{r}_1 - \boldsymbol{\rho}_S)\Theta(-\mathbf{v}_1 \cdot \hat{\mathbf{n}})f(\mathbf{r}_1, \mathbf{v}_1, t). \quad (28)$$

The determination of an expression for  $\Gamma_W^{(+)}$  is very similar to that just given for the loss term. However we do need to make some further assumptions in order to obtain an expression for the gain term. These assumptions are

1. If a particle collides with the surface, it leaves the surface instantly at the point where the impact takes place.
2. When a particle with velocity  $\mathbf{v}'_1$  collides with the surface, the probability that it leaves the surface with velocity  $\mathbf{v}_1$  in the range  $\delta\mathbf{v}_1$  is given by  $P(\mathbf{v}_1, \mathbf{v}'_1)\delta\mathbf{v}_1$ , where  $P$  is taken to be independent of the time,  $t$ , and to satisfy the following conditions:
  - (a)  $P(\mathbf{v}_1, \mathbf{v}'_1)$  vanishes unless  $\mathbf{v}'_1 \cdot \hat{\mathbf{n}} < 0$  and  $\mathbf{v}_1 \cdot \hat{\mathbf{n}} > 0$ , so that  $P$  can be expressed in terms of a collision kernel  $K_W(\mathbf{v}_1, \mathbf{v}'_1)$  as

$$P(\mathbf{v}_1, \mathbf{v}'_1) = K_W(\mathbf{v}_1, \mathbf{v}'_1)\Theta(-\mathbf{v}'_1 \cdot \hat{\mathbf{n}})\Theta(\mathbf{v}_1 \cdot \hat{\mathbf{n}}). \quad (29)$$

- (b)  $P(\mathbf{v}'_1, \mathbf{v}_1)$  is properly normalized, so that

$$\int_{\mathbf{v}_1 \cdot \hat{\mathbf{n}} > 0} d\mathbf{v}_1 P(\mathbf{v}_1, \mathbf{v}'_1) = 1 \text{ for } \mathbf{v}'_1 \cdot \hat{\mathbf{n}} < 0. \quad (30)$$

- (c) Since we have assumed that particles striking the boundaries are instantly reflected from the point of impact in a way that is described by the collision kernel  $K_W$ , we must impose the condition that the rate at which particles leave the surface at the point of impact should be equal to the rate at which particles arrive at that point on the surface. This condition is expressed by the relation

$$\begin{aligned} |\mathbf{v}_1 \cdot \hat{\mathbf{n}}|\Theta(\mathbf{v}_1 \cdot \hat{\mathbf{n}})f(\boldsymbol{\rho}_S, \mathbf{v}_1, t) &= \\ &= \int_{\mathbf{v}'_1 \cdot \hat{\mathbf{n}} < 0} d\mathbf{v}'_1 |\mathbf{v}'_1 \cdot \hat{\mathbf{n}}| K_W(\mathbf{v}_1, \mathbf{v}'_1) f(\boldsymbol{\rho}_S, \mathbf{v}'_1, t). \end{aligned} \quad (31)$$

The precise specification of the collision kernel,  $K_W$  depends upon the nature of the interactions of the gas particles with the wall. Moreover, one might also consider relaxing the conditions expressed by Eqs. (30) and (31) to treat cases where particles are gained or lost at the boundaries, if part of the boundary is a droplet, for example. Here we will avoid specifying the kernel

as much as possible, but always assume that the normalization condition, Eq. (30) is satisfied. Using these assumptions, we can easily construct the following expression for  $\Gamma_W^{(+)}$ , the rate at which particles with velocity  $\mathbf{v}_1$  are created due to collisions with the walls:

$$\Gamma_W^{(+)} = \int_{\partial V} dS \delta(\mathbf{r}_1 - \boldsymbol{\rho}_S) \int d\mathbf{v}'_1 |\mathbf{v}'_1 \cdot \hat{\mathbf{n}}| \times \Theta(-\mathbf{v}'_1 \cdot \hat{\mathbf{n}}) \Theta(\mathbf{v}_1 \cdot \hat{\mathbf{n}}) K_W(\mathbf{v}_1, \mathbf{v}'_1) f(\mathbf{r}_1, \mathbf{v}'_1, t). \quad (32)$$

Equations (28) and (33) can be combined to obtain the expression for  $\Gamma_W$ , which is

$$\Gamma_W = \int_{\partial V} dS \delta(\mathbf{r}_1 - \boldsymbol{\rho}_S) \left[ \Theta(\mathbf{v}_1 \cdot \hat{\mathbf{n}}) \int d\mathbf{v}'_1 \Theta(-\mathbf{v}'_1 \cdot \hat{\mathbf{n}}) |\mathbf{v}'_1 \cdot \hat{\mathbf{n}}| K_W(\mathbf{v}_1, \mathbf{v}'_1) f(\mathbf{r}_1, \mathbf{v}'_1, t) - \Theta(-\mathbf{v}_1 \cdot \hat{\mathbf{n}}) |\mathbf{v}_1 \cdot \hat{\mathbf{n}}| f(\mathbf{r}_1, \mathbf{v}_1, t) \right]. \quad (33)$$

We take the opportunity to point out already here that the expression for  $\Gamma_W$  just derived has the property that it keeps particles inside the container. That is to say the solutions of the Boltzmann equation automatically have the property that they vanish for spatial coordinates that are beyond the inner surface of the vessel containing the gas, if they did so at the initial time. Mathematically, the solutions of interest to us can be expressed in the form

$$f(\mathbf{r}, \mathbf{v}, t) = W_W(\mathbf{r}) \tilde{f}(\mathbf{r}, \mathbf{v}, t), \quad (34)$$

where  $W_W(\mathbf{r})$  vanishes whenever  $\mathbf{r}$  is located outside the bounding surface of the gas, and is unity otherwise. The distribution function  $\tilde{f}$  is identical to  $f$  inside the vessel, but may be continued outside of it so as to be continuous across the boundary of the vessel, while  $f$  is not continuous under any circumstances.

It will turn out convenient to rewrite expression Eq. (33) for  $\Gamma_W$  in terms of a *wall collision operator*,  $\bar{\mathbf{T}}_{W-}$ , in an obvious notation, as  $\mathbf{T}_W f$  as

$$\Gamma_W \equiv \bar{\mathbf{T}}_{W-} f(\mathbf{r}_1, \mathbf{v}_1, t), \quad (35)$$

where the wall collision operator,  $\bar{\mathbf{T}}_{W-}$  is given by

$$\begin{aligned} \bar{\mathbf{T}}_{W-} &= \int_{\partial V} dS \delta(\mathbf{r}_1 - \boldsymbol{\rho}_S) \\ &\times \left[ \Theta(\mathbf{v}_1 \cdot \hat{\mathbf{n}}) \int d\mathbf{v}'_1 \Theta(-\mathbf{v}'_1 \cdot \hat{\mathbf{n}}) |\mathbf{v}'_1 \cdot \hat{\mathbf{n}}| K_W(\mathbf{v}_1, \mathbf{v}'_1) \mathbf{b} \right. \\ &\left. - \Theta(-\mathbf{v}_1 \cdot \hat{\mathbf{n}}) |\mathbf{v}_1 \cdot \hat{\mathbf{n}}| \right]. \end{aligned} \quad (36)$$

Here  $\mathbf{b}$  is an exchange operator which replaces the velocity  $\mathbf{v}_1$  by  $\mathbf{v}'_1$  in all functions of  $\mathbf{v}_1$  appearing to it's right in any expression in which it appears.

## 1.5 The Boltzmann Transport Equation

We have now completed the derivation of the Boltzmann transport equation and established a notation that will be useful in our further discussions, both for the Boltzmann equation and for the generalized Boltzmann equation needed to treat dense gases. Combining the results of the previous sections we can write the Boltzmann transport equation as

$$\begin{aligned} & \frac{\partial f(\mathbf{r}_1, \mathbf{v}_1, t)}{\partial t} + \nabla_{\mathbf{r}_1} \cdot (\mathbf{v} f(\mathbf{r}_1, \mathbf{v}_1, t)) + \nabla_{\mathbf{v}} \cdot (\dot{\mathbf{v}} f(\mathbf{r}_1, \mathbf{v}_1, t)) = \\ & = \int d\mathbf{v}_2 \int d\hat{\sigma} B(\mathbf{g}, \hat{\sigma})(\mathbf{b}_{\hat{\sigma}}(1, 2) - 1) f(\mathbf{r}_1, \mathbf{v}_1, t) f(\mathbf{r}_2, \mathbf{v}_2, t) + \bar{\mathbf{T}}_W - f(\mathbf{r}_1, \mathbf{v}_1) \end{aligned} \quad (37)$$

## 2 The $H$ -Theorem of Boltzmann

It follows from experiment and from equilibrium statistical mechanics that for a dilute gas confined to a container at rest and in thermodynamic equilibrium in an external field  $\phi_{ext}(\mathbf{r})$ , the single particle distribution function,  $f_{eq}(\mathbf{r}, \mathbf{v})$  has the Maxwell-Boltzmann form

$$f(\mathbf{r}, \mathbf{v}) = n W_W(\mathbf{r}) \left( \frac{\beta m}{2\pi} \right)^{\frac{3}{2}} C e^{-\beta[\frac{1}{2}m\mathbf{v}^2 + \phi_{ext}(\mathbf{r})]} \quad (38)$$

where  $n$  is the number density of the gas,  $n = N/V$ ,  $\beta = (k_B T)^{-1}$ , where  $k_B$  is Boltzmann's constant,  $T$  is the thermodynamic temperature,  $W_W(\mathbf{r})$  is the step function defined below Eq. (34), and  $C$  is a normalization constant given by

$$C = V \left[ \int_{all\ space} d\mathbf{r} W_W(\mathbf{r}) e^{-\beta\phi_{ext}(\mathbf{r})} \right]^{-1}. \quad (39)$$

It is therefore natural to ask whether or not this result is in fact predicted by the Boltzmann equation. That is to say:

1. Can one show, using the Boltzmann equation, that a dilute gas not in equilibrium at some time will, as time progresses, approach an equilibrium state with distribution function given by Eq. (38), and that once in this state, the gas will remain in this state?

2. Can one describe in detail how the approach to this equilibrium state proceeds?

Boltzmann showed that one can give an affirmative answer to the first question without having to construct a solution to the equation, assuming of course, that a solution to the equation exists and that it satisfies some smoothness and integrability conditions. To do this he proved a theorem, the *Boltzmann H-Theorem* which states that the function  $H(t)$  defined by

$$H(t) = \int_V d\mathbf{r}d\mathbf{v} f(\mathbf{r}, \mathbf{v}, t) (\ln f(\mathbf{r}, \mathbf{v}, t) - 1), \quad (40)$$

is a monotonically decreasing function of time provided that  $f(\mathbf{r}, \mathbf{v}, t)$  is a solution and there is no energy exchange between the gas and the outside world, including the walls of the vessel<sup>4</sup> The form of the  $H$  function was suggested by the fact that if one replaces  $f$  in Eq.(40) by  $f_{eq}$  given by Eq. (38), then  $H = -S/k_B + constant$ , where  $S$  is the equilibrium entropy of an ideal gas of  $N$  particles in equilibrium in volume  $V$  at temperature  $T$ . The  $H$  function ceases to change with time when the distribution function takes the equilibrium, Maxwell-Boltzmann form as given by Eq. (38). In order to check this let us substitute this form into the Boltzmann equation (37). One easily checks that the second and third term on the left hand side of this equation cancel each other. Furthermore, the first term on the right hand side,  $J(f, f)$ , also vanishes as a consequence of the energy conservation during a collision as expressed in Eq. (8). So the Maxwell-Boltzmann distribution will be stationary indeed provided the action of  $\bar{\mathbf{T}}_{W-}$  upon it will vanish if the wall is in equilibrium with the gas. A sufficient condition for this is the so-called *thermostat condition*, which is of the form

$$\begin{aligned} & \Theta(\mathbf{v} \cdot \hat{\mathbf{n}}) |\mathbf{v} \cdot \hat{\mathbf{n}}| \phi_W(\mathbf{v}, \rho_S) = \\ & = \Theta(\mathbf{v} \cdot \hat{\mathbf{n}}) \int d\mathbf{v}' \Theta(-\mathbf{v}' \cdot \hat{\mathbf{n}}) |\mathbf{v}' \cdot \hat{\mathbf{n}}| K_W(\mathbf{v}, \mathbf{v}') \phi_W(\mathbf{v}', \rho_S), \end{aligned} \quad (41)$$

---

<sup>4</sup>Boltzmann in his derivation of the H-theorem did not explicitly take into account any interactions between gas particles and the walls, but his treatment does allow for walls that can be represented by fixed external potentials. Here we will present a derivation of the H-theorem that allows for more general ways of representing the interactions with the walls. Still we will have to impose the so-called thermostat condition on these interactions to be able to derive the H-theorem.

where  $\phi_W(\mathbf{v}, \rho_S)$  is a Maxwell-Boltzmann distribution with temperature  $T_W(\rho_S)$  given by <sup>5</sup>

$$\phi_W(\mathbf{v}, \rho_S) = \left( \frac{m}{2\pi k_B T_W(\rho_S)} \right)^{3/2} e^{-\left( \frac{m\mathbf{v}^2}{2k_B T_W(\rho_S)} \right)}. \quad (42)$$

In the sequel we will assume throughout that this thermostat condition is satisfied.

The  $H$ -Theorem naturally provoked a considerable amount of controversy at the time of its publication, in part because of the inherent conflict between the monotonic decrease of Boltzmann's  $H$  with time and the time reversibility of Newton's equations of motion, and in part because the atomic theory of matter had not at the time (1872) been firmly established. Here we present the  $H$ -Theorem in a form appropriate for a gas in a bounded system that allows an exchange of energy between the gas and the boundaries. We will then discuss in some detail the source of irreversibility in the solutions of the Boltzmann equation, explain how one can resolve the two famous paradoxes posed by Zermelo and Loschmidt and argue why the Boltzmann equation cannot be a direct consequence of Newton's equations of mechanics. Nevertheless, we will argue that the  $H$ -Theorem provides, for dilute gases at least, a statistical mechanical framework for understanding the molecular origins of the Second Law of Thermodynamics<sup>6</sup>.

## 2.1 Some Models for the Interactions of the Gas Particles with the Boundaries

The interactions between gas particles and the physical boundaries of the vessel containing them can be quite complicated. There is a considerable literature concerning the exchange of energy between the gas and the boundaries and many theoretical treatments of it. Since our focus is not on the details of the physical processes responsible for this exchange, we will use here simplified models that capture some of the important features necessary for understanding Boltzmann's  $H$ -Theorem.

The simplest type of gas-surface interaction is described by specular, elas-

---

<sup>5</sup>We continue to use the case of three dimensional systems as an example. All of the expressions given here can easily be generalized to any number of spatial dimensions.

<sup>6</sup>Boltzmann wrote in his book *Lectures on Gas Theory* that "... this proof has, as it will appear, a not uninteresting connection with the entropy principle,..."

tic collisions of the gas molecules with a surface at rest<sup>7</sup>, as would result from a hard-core wall potential which is infinite outside the vessel and vanishes inside<sup>8</sup>. In such a case the collision kernel of Eq. (29) is given by

$$K_W^{sp}(\mathbf{v}_1, \mathbf{v}'_1) = \delta(\mathbf{v}'_1 - \mathbf{v}_1 + 2\hat{\mathbf{n}}(\hat{\mathbf{n}} \cdot \mathbf{v}_1)). \quad (43)$$

Here  $\hat{\mathbf{n}}$  is a unit vector normal to the boundary and chosen to point into the interior of the gas. Notice that this reflection law trivially satisfies the thermostat condition.

Although this is a simple boundary condition, which will be used often in the coming sections, it is by no means the only boundary condition that one might naturally think of. A very natural and useful boundary condition is constructed by supposing that particles striking the boundary at a point  $\boldsymbol{\rho}_S$  are re-emitted with a Maxwell-Boltzmann velocity distribution with a temperature,  $T_W(\boldsymbol{\rho}_S)$  characteristic of the wall at the point of incidence. From the thermostat condition (41) together with the normalization condition (35) one readily finds that the *thermal reflection* kernel is given by

$$K_W^{th}(\mathbf{v}, \mathbf{v}') = \left( \frac{mk_B T_w(\boldsymbol{\rho}_S)}{2\pi} \right)^{-1/2} \Theta(\mathbf{v} \cdot \hat{\mathbf{n}}) |\mathbf{v} \cdot \hat{\mathbf{n}}| \phi_W(\mathbf{v}, \boldsymbol{\rho}_S). \quad (44)$$

## 2.2 Proof of the $H$ -Theorem

We now turn to the proof of the  $H$ -Theorem in a form that is appropriate under the thermostat condition discussed above. That is we will prove that if the kernel  $K_W(\mathbf{v}, \mathbf{v}')$  satisfies the thermostat condition, then the time rate of change of the  $H$  function satisfies

$$\frac{dH}{dt} \leq -\frac{1}{k_B} \int_{\partial V} dS \frac{1}{T_W(\boldsymbol{\rho}_S)} \frac{dQ(\boldsymbol{\rho}_S)}{dt}, \quad (45)$$

---

<sup>7</sup>Reflection laws for moving surfaces are found most easily by first formulating them in the rest frame of the surface and then transforming to the rest frame of the gas

<sup>8</sup>For smooth wall potentials which increase from zero to infinity over a small distance, the specular collision kernel is not exact, but it is an excellent approximation, comparable to that of the neglect of the difference in position of two colliding particles. The effect of such potentials may also be accounted for by including them in the external potential determining  $\dot{\mathbf{v}}$  in Eq. (37). In most cases however, the representation by the specular collision kernel, or equivalently by a matching condition between the distribution functions with incoming and reflected velocities, is simpler

where the integration is over the entire bounding surface of the gas,  $dQ(\boldsymbol{\rho}_S)/dt$  is the energy flux from the walls to the gas at the point  $\boldsymbol{\rho}_S$  on the boundary and  $T(\boldsymbol{\rho}_S)$  is the temperature at that point. We hold the boundary fixed so that there is no work done on the gas, and the energy flow is equal to the heat flow. The net heat flow into the gas will be given by

$$\frac{dQ(\boldsymbol{\rho}_S)}{dt} = \int d\mathbf{v} (\mathbf{v} \cdot \hat{\mathbf{n}}) \frac{m\mathbf{v}^2}{2} f(\boldsymbol{\rho}_S, \mathbf{v}, t). \quad (46)$$

Since the integration is over all values of  $\mathbf{v}$ , the heat flow includes both the energy transferred to the walls and that transferred to the gas.

Using the definition of the  $H$  function given by Eq. (40), we define a density of  $H(t)$  by

$$h(\mathbf{r}, t) = \int d\mathbf{v} f(\mathbf{r}, \mathbf{v}, t) [\ln f(\mathbf{r}, \mathbf{v}, t) - 1], \quad (47)$$

indeed satisfying

$$H(t) = \int d\mathbf{r} h(\mathbf{r}, t). \quad (48)$$

It then follows from Eq. (24) that

$$\begin{aligned} \frac{\partial h(\mathbf{r}, t)}{\partial t} &= \int d\mathbf{v} \frac{\partial f(\mathbf{r}, \mathbf{v}, t)}{\partial t} \ln f(\mathbf{r}, \mathbf{v}, t) \\ &= -\nabla_{\mathbf{r}} \cdot \mathbf{j}_h \\ &+ \int d\mathbf{v} \ln f(\mathbf{r}, \mathbf{v}, t) \left\{ \int d\mathbf{v}_1 \int d\hat{\boldsymbol{\sigma}} B(\mathbf{g}, \hat{\boldsymbol{\sigma}}) [f' f'_1 - f f_1] \right\} \\ &+ \int d\mathbf{v} \ln f(\mathbf{r}, \mathbf{v}, t) \mathbf{T}_W f(\mathbf{r}, \mathbf{v}, t). \end{aligned} \quad (49)$$

We have used the notation introduced in Eq. (25) and simplified the notation so that  $(\mathbf{r}_1, \mathbf{v}_1) \rightarrow (\mathbf{r}, \mathbf{v})$  and  $(\mathbf{r}_2, \mathbf{v}_2) \rightarrow (\mathbf{r}_1, \mathbf{v}_1)$ . Here  $\mathbf{j}_h$  is the flux in  $h$  due to the free motion of the particles. To obtain Eq. (49) we have used that for physical systems  $\nabla_{\mathbf{v}} \cdot \dot{\mathbf{v}} = 0$  and assumed that  $f(\mathbf{r}, \mathbf{v}, t)$  vanishes sufficiently rapidly at large velocities so that

$$\int d\mathbf{v} \nabla_{\mathbf{v}} \cdot [\dot{\mathbf{v}}(f \ln f - f)] \rightarrow 0. \quad (50)$$

This will certainly be the case if the total number of particles and the total kinetic energy of the particles in volume  $V$  are both bounded.

Next we consider the change in  $h$  due to binary collisions, the third term on the right hand side of Eq. (49). To analyze this term we first establish a more general and very useful identity:

$$\begin{aligned} \int d\mathbf{v}\psi(\mathbf{v})J(f, f) &= \frac{1}{4} \int d\mathbf{v} \int d\mathbf{v}_1 \int d\hat{\boldsymbol{\sigma}} B(\mathbf{g}, \hat{\boldsymbol{\sigma}}) (f'f'_1 - ff_1) \\ &\times [\psi(\mathbf{v}) + \psi(\mathbf{v}_1) - \psi(\mathbf{v}') - \psi(\mathbf{v}'_1)]. \end{aligned} \quad (51)$$

We start by noting that

$$\begin{aligned} \int d\mathbf{v}\psi(\mathbf{v})J(f, f) &= \frac{1}{2} \int d\mathbf{v} \int d\mathbf{v}_1 \int d\hat{\boldsymbol{\sigma}} B(\mathbf{g}, \hat{\boldsymbol{\sigma}}) [\psi(\mathbf{v}) + \psi(\mathbf{v}_1)] \\ &\times (f'f'_1 - ff_1). \end{aligned} \quad (52)$$

Here we just exchanged the variables  $\mathbf{v}$  and  $\mathbf{v}_1$ , summed and divided the sum by two. Next we replace  $B(\mathbf{g}, \hat{\boldsymbol{\sigma}})$  by  $B(\mathbf{g}', -\hat{\boldsymbol{\sigma}})$ ;  $\hat{\boldsymbol{\sigma}}$  by  $-\hat{\boldsymbol{\sigma}}$  and  $d\mathbf{v}d\mathbf{v}_1$  by  $d\mathbf{v}'d\mathbf{v}'_1$ , to write

$$\begin{aligned} \int d\mathbf{v}\psi(\mathbf{v})J(f, f) &= \frac{1}{2} \int d\hat{\boldsymbol{\sigma}} \int d\mathbf{v}' \int d\mathbf{v}'_1 B(\mathbf{g}', -\hat{\boldsymbol{\sigma}}) [\psi(\mathbf{v}) + \psi(\mathbf{v}_1)] \\ &\times (f'f'_1 - ff_1). \end{aligned} \quad (53)$$

At this point one does something interesting: We take advantage of the fact that the velocities,  $(\mathbf{v}', \mathbf{v}'_1)$  are the initial velocities in the restituting collision in which  $(\mathbf{v}, \mathbf{v}_1)$  are the final velocities. This allows us to remove the primes in the integration variables on the right hand side of Eq. (53). Next we can exchange primed and unprimed variables in this equation to obtain

$$\begin{aligned} \int d\mathbf{v}\psi(\mathbf{v})J(f, f) &= \frac{1}{2} \int d\hat{\boldsymbol{\sigma}} \int d\mathbf{v} \int d\mathbf{v}_1 B(\mathbf{g}, \hat{\boldsymbol{\sigma}}) [\psi(\mathbf{v}') + \psi(\mathbf{v}'_1)] \\ &\times (ff_1 - f'f'_1). \end{aligned} \quad (54)$$

If we now combine Eqs. (53) and (54) and divide again by two, we obtain Eq. (51).

Applying this result to Eq. (49) we now obtain

$$\begin{aligned}
\frac{\partial h(\mathbf{r}, t)}{\partial t} &= -\nabla \cdot \mathbf{j}_h + \\
&+ \frac{1}{4} \int d\mathbf{v} \int d\mathbf{v}_1 \int d\hat{\boldsymbol{\sigma}} B(\mathbf{g}, \hat{\boldsymbol{\sigma}}) [f' f'_1 - f f_1] \ln \frac{f f_1}{f' f'_1} + \\
&+ \int d\mathbf{v} \mathbf{T}_W f \ln f. \tag{55}
\end{aligned}$$

The second term on the right hand side is either negative or zero, since  $(b-a) \ln(a/b) \leq 0$  for any real positive  $a$  and  $b$ . The equality sign only holds if  $a = b$ , of course. As we will see below, the equality  $f f_1 = f' f'_1$  leads to the Maxwell-Boltzmann distribution. For the moment we continue with the derivation of the  $H$ -Theorem, which we can now write in the form

$$\frac{\partial h(\mathbf{r}, t)}{\partial t} \leq -\nabla \cdot \mathbf{j}_h + \int d\mathbf{v} \ln f \bar{\mathbf{T}}_W - f. \tag{56}$$

To obtain an inequality for  $H(t)$  we integrate over all space, use the fact that the distribution function vanishes outside of the container and find that

$$\begin{aligned}
\frac{dH(t)}{dt} &\leq \int d\mathbf{r} \int d\mathbf{v} \ln f \bar{\mathbf{T}}_W - f, \text{ or} \\
\frac{dH(t)}{dt} &\leq \int_{\partial V} dS \int d\mathbf{v} \ln f(\boldsymbol{\rho}_S, \mathbf{v}, t) \times \\
&\times \left\{ \Theta(\mathbf{v} \cdot \hat{\mathbf{n}}) \int d\mathbf{v}' \Theta(-\mathbf{v}' \cdot \hat{\mathbf{n}}) |\mathbf{v}' \cdot \hat{\mathbf{n}}| K_W(\mathbf{v}, \mathbf{v}') f(\boldsymbol{\rho}_S, \mathbf{v}', t) \right. \\
&\left. - |\mathbf{v} \cdot \hat{\mathbf{n}}| f(\boldsymbol{\rho}_S, \mathbf{v}, t) \Theta(-\mathbf{v} \cdot \hat{\mathbf{n}}) \right\}. \tag{57}
\end{aligned}$$

The last equation can be transformed to a form that will be convenient for the application of another mathematical inequality, Jensen's inequality, if we use the definition of the wall collision kernel,  $K_W$  and the normalization condition given by Eqs. (29) and (30), respectively, to write

$$\begin{aligned}
\frac{dH(t)}{dt} &\leq \int_{\partial V} dS \int d\mathbf{v} \int d\mathbf{v}' \Theta(\mathbf{v} \cdot \hat{\mathbf{n}}) \Theta(-\mathbf{v}' \cdot \hat{\mathbf{n}}) |\mathbf{v}' \cdot \hat{\mathbf{n}}| K_W(\mathbf{v}, \mathbf{v}') \times \\
&\times [f(\boldsymbol{\rho}_S, \mathbf{v}', t) \ln f(\boldsymbol{\rho}_S, \mathbf{v}, t) - f(\boldsymbol{\rho}_S, \mathbf{v}, t) \ln f(\boldsymbol{\rho}_S, \mathbf{v}', t)]. \tag{58}
\end{aligned}$$

We now wish to make use of the thermostat condition, Eq. (41) and therefore we introduce the Maxwell-Boltzmann distribution function,  $\phi_W$  defined by

Eq. (42) into the previous equation by writing

$$f(\boldsymbol{\rho}_S, \mathbf{v}, t) \equiv \phi_W(\boldsymbol{\rho}_S, \mathbf{v}) \hat{f}(\boldsymbol{\rho}_S, \mathbf{v}, t), \quad (59)$$

so as to rewrite Eq. (58) as

$$\begin{aligned} \frac{dH(t)}{dt} &\leq \int_{\partial V} dS \int d\mathbf{v} \int d\mathbf{v}' \Theta(\mathbf{v} \cdot \hat{\mathbf{n}}) \Theta(-\mathbf{v}' \cdot \hat{\mathbf{n}}) |\mathbf{v}' \cdot \hat{\mathbf{n}}| K_W(\mathbf{v}, \mathbf{v}') \phi_W(\boldsymbol{\rho}_S, \mathbf{v}') \\ &\times \left[ \hat{f}(\boldsymbol{\rho}_S, \mathbf{v}', t) \ln \hat{f}(\boldsymbol{\rho}_S, \mathbf{v}, t) - \hat{f}(\boldsymbol{\rho}_S, \mathbf{v}', t) \ln \hat{f}(\boldsymbol{\rho}_S, \mathbf{v}', t) + \right. \\ &\left. + \hat{f}(\boldsymbol{\rho}_S, \mathbf{v}', t) \ln \phi_W(\boldsymbol{\rho}_S, \mathbf{v}) - \hat{f}(\boldsymbol{\rho}_S, \mathbf{v}', t) \ln \phi_W(\boldsymbol{\rho}_S, \mathbf{v}') \right]. \end{aligned} \quad (60)$$

The crucial step in the argument now is to use the fact that the function  $x \ln x$  is a convex function for all  $x \geq 0$ . That is, the straight line connecting any two points  $p_1$  and  $p_2$  on the curve always lies above the curve. Expressed mathematically, any such convex function  $\psi(x)$  satisfies the inequality,

$$\psi\left(\sum_i a_i x_i\right) \leq \sum_i a_i \psi(x_i), \quad (61)$$

where the  $a_i$  are real, positive numbers, satisfying  $\sum_i a_i = 1$ . Here we use an integral form of this inequality, known as Jensen's inequality, which we express in the form needed here as

$$\frac{\int dx f(x) g(x)}{\int dx g(x)} \ln \left[ \frac{\int dx f(x) g(x)}{\int dx g(x)} \right] \leq \frac{\int dx g(x) f(x) \ln f(x)}{\int dx g(x)}, \quad (62)$$

where  $f(x)$  and  $g(x)$  are real, positive functions.

To apply Jensen's inequality to Eq. (60), we set the region of integration to be over all  $\mathbf{v}'$  with  $\mathbf{v}' \cdot \hat{\mathbf{n}} < 0$  and let

$$\begin{aligned} g(\mathbf{v}') &= |\mathbf{v}' \cdot \hat{\mathbf{n}}| \Theta(\mathbf{v} \cdot \hat{\mathbf{n}}) K_W(\mathbf{v}, \mathbf{v}') \phi_W(\boldsymbol{\rho}_S, \mathbf{v}') \\ f(\mathbf{v}') &= \hat{f}(\boldsymbol{\rho}_S, \mathbf{v}', t). \end{aligned}$$

When Jensen's inequality is applied to these functions we obtain

$$\begin{aligned} \Theta(\mathbf{v} \cdot \hat{\mathbf{n}}) |\mathbf{v} \cdot \hat{\mathbf{n}}| f(\boldsymbol{\rho}_S, \mathbf{v}, t) \ln \hat{f}(\boldsymbol{\rho}_S, \mathbf{v}, t) \int d\mathbf{v}' \Theta(-\mathbf{v}' \cdot \hat{\mathbf{n}}) |\mathbf{v}' \cdot \hat{\mathbf{n}}| \leq \\ \Theta(\mathbf{v} \cdot \hat{\mathbf{n}}) K_W(\mathbf{v}, \mathbf{v}') f(\boldsymbol{\rho}_S, \mathbf{v}', t) \ln \hat{f}(\boldsymbol{\rho}_S, \mathbf{v}', t). \end{aligned} \quad (63)$$

Here we used Eq. (31) and the thermostat condition, Eq. (41), as well as the identity

$$\Theta(\mathbf{v} \cdot \hat{\mathbf{n}}) \ln \left[ \frac{|\mathbf{v} \cdot \hat{\mathbf{n}}| \Theta(\mathbf{v} \cdot \hat{\mathbf{n}}) f(\boldsymbol{\rho}_S, \mathbf{v}, t)}{|\mathbf{v} \cdot \hat{\mathbf{n}}| \Theta(\mathbf{v} \cdot \hat{\mathbf{n}}) \phi_W(\boldsymbol{\rho}_S, \mathbf{v})} \right] = \Theta(\mathbf{v} \cdot \hat{\mathbf{n}}) \ln \hat{f}(\boldsymbol{\rho}_S, \mathbf{v}, t). \quad (64)$$

The inequality given by Eq. (62) allows us to strengthen the inequality for  $dH/dt$ , Eq. (60), by subtracting a smaller quantity for the right hand side of this equation. That is, we find that

$$\begin{aligned} \frac{dH(t)}{dt} &\leq \int_{\partial V} dS \int d\mathbf{v} \int d\mathbf{v}' \Theta(\mathbf{v} \cdot \hat{\mathbf{n}}) \Theta(-\mathbf{v}' \cdot \hat{\mathbf{n}}) |\mathbf{v}' \cdot \hat{\mathbf{n}}| K_W(\mathbf{v}, \mathbf{v}') \phi_W(\boldsymbol{\rho}_S, \mathbf{v}') \times \\ &\times \left[ \hat{f}(\boldsymbol{\rho}_S, \mathbf{v}', t) \ln \hat{f}(\boldsymbol{\rho}_S, \mathbf{v}, t) + \right. \\ &+ \left. \hat{f}(\boldsymbol{\rho}_S, \mathbf{v}', t) \ln \phi_W(\boldsymbol{\rho}_S, \mathbf{v}) - \hat{f}(\boldsymbol{\rho}_S, \mathbf{v}', t) \ln \phi_W(\boldsymbol{\rho}_S, \mathbf{v}') \right] \\ &- \int_{\partial V} \int d\mathbf{v} \Theta(\mathbf{v} \cdot \hat{\mathbf{n}}) |\mathbf{v} \cdot \hat{\mathbf{n}}| f(\boldsymbol{\rho}_S, \mathbf{v}, t) \ln \hat{f}(\boldsymbol{\rho}_S, \mathbf{v}, t). \end{aligned} \quad (65)$$

One more application of Eq. (31) shows that the first and last terms on the right hand side of Eq. (65) cancel and we are left with

$$\begin{aligned} \frac{dH(t)}{dt} &\leq \int_{\partial V} dS \int d\mathbf{v} \int d\mathbf{v}' \Theta(\mathbf{v} \cdot \hat{\mathbf{n}}) \Theta(-\mathbf{v}' \cdot \hat{\mathbf{n}}) |\mathbf{v}' \cdot \hat{\mathbf{n}}| K_W(\mathbf{v}, \mathbf{v}') \phi_W(\boldsymbol{\rho}_S, \mathbf{v}') \\ &\times \left[ \hat{f}(\boldsymbol{\rho}_S, \mathbf{v}', t) \ln \phi_W(\boldsymbol{\rho}_S, \mathbf{v}) - \hat{f}(\boldsymbol{\rho}_S, \mathbf{v}', t) \ln \phi_W(\boldsymbol{\rho}_S, \mathbf{v}') \right]. \end{aligned} \quad (66)$$

If we now insert the expression for  $\phi_W(\boldsymbol{\rho}_S, \mathbf{v})$ , Eq. (42), in Eq. (66), use Eqs. (30) and (31), we obtain

$$\frac{dH(t)}{dt} \leq -\frac{1}{k_B} \int_{\partial V} dS \frac{1}{T_W(\boldsymbol{\rho}_S)} \int d\mathbf{v} (\mathbf{v} \cdot \hat{\mathbf{n}}) \frac{m\mathbf{v}^2}{2} f(\boldsymbol{\rho}_S, \mathbf{v}, t), \quad (67)$$

where the velocity integration is over all  $\mathbf{v}$ . This is the form of the  $H$ -theorem announced by Eqs. (45) and (46).

The original form of Boltzmann's  $H$ -theorem,  $dH/dt \leq 0$ , is recovered if there is no net energy flow into or out of the container. This can occur if the particles make elastic, specular collisions with the walls, or if the wall temperature is uniform, and the velocity distribution function of the gas is a Maxwell-Boltzmann distribution with the temperature equal to that of the walls.

## 2.3 The Equilibrium Distribution Function

Now we turn our attention to Boltzmann's argument that the  $H$ -Theorem can be used to show that the gas will eventually reach a state of total equilibrium with temperature equal to the wall temperature, provided the particle-wall interactions satisfy the thermostat condition, Eq. (41), and the wall temperature has a uniform value  $T_W$ . Our goal is to show that under these circumstances the distribution function takes the form given by Eq. (38).

We begin by noting that we can define a time dependent, non-equilibrium Helmholtz free energy,  $F(t)$  by<sup>9</sup>

$$F(t) = E(t) - T_W S(t) = E(t) + k_B T_W H(t), \quad (68)$$

where the energy function,  $E(t)$  is given by

$$E(t) = \int d\mathbf{r} \int d\mathbf{v} \left[ \frac{m\mathbf{v}^2}{2} + \phi_{ext}(\mathbf{r}) \right] f(\mathbf{r}, \mathbf{v}, t). \quad (69)$$

It follows from the Boltzmann equation and the  $H$ -Theorem that

$$\frac{dF(t)}{dt} \leq 0. \quad (70)$$

This is a well-known expression of Clausius' inequality applied to systems in contact with a heat bath, except that, for a dilute gas the definition of the free energy has been generalized towards arbitrary non-equilibrium states. If we assume that the total energy,  $E(t)$  is bounded from below and the total number of particles,  $N$ , is fixed we easily find that the Helmholtz free energy is also bounded from below<sup>10</sup>. Hence from Eq. (69) it follows that  $F(t)$  is a bounded, monotonic non-increasing function of time. It then approaches a limit as  $t \rightarrow \infty$  such that  $dF(t)/dt \rightarrow 0$ , in this limit. The observation that the Helmholtz free energy eventually approaches a constant value is the essence of the argument for the approach of the distribution function to an

---

<sup>9</sup>We have assumed that the walls of the container are kept fixed. We can relax this condition; e.g. for systems with mobile walls kept under constant external pressure we could use the Gibbs free energy instead of the Helmholtz free energy in the argument presented below Eq. (68.)

<sup>10</sup>Variational calculus applied to Eq. (69) shows that the minimum is obtained for a Maxwell-Boltzmann distribution with the temperature  $T_W$ .

equilibrium form. An examination of the proof of the  $H$ -Theorem shows that  $dF(t)/dt = 0$  if and only if the equality

$$f(\mathbf{r}, \mathbf{v}, t)f(\mathbf{r}, \mathbf{v}_1, t) = f(\mathbf{r}, \mathbf{v}', t)f(\mathbf{r}, \mathbf{v}'_1, t) \quad (71)$$

is satisfied. Taking logarithms we find that Eq. (71) is equivalent to the statement that  $\ln f$  is conserved in a binary collision. That is

$$\ln f(\mathbf{r}, \mathbf{v}, t) + \ln f(\mathbf{r}, \mathbf{v}_1, t) = \ln f(\mathbf{r}, \mathbf{v}', t) + \ln f(\mathbf{r}, \mathbf{v}'_1, t). \quad (72)$$

From this result it follows that the logarithm of the distribution function satisfying Eq. (71) must be of the form

$$\ln f(\mathbf{r}, \mathbf{v}, t) = a(\mathbf{r}, t) + \mathbf{b}(\mathbf{r}, t) \cdot \mathbf{v} + c(\mathbf{r}, t)\mathbf{v}^2, \quad (73)$$

since the quantities conserved in a binary collision are the number of particles, the total momentum and the total energy of the system of two particles. One might also include the total angular momentum about some point, but this would be redundant. This follows from the fact that all four distribution functions in Eq. (72) are evaluated at the same point,  $\mathbf{r}$ , so the conservation of angular momentum,  $\mathbf{r} \times \mathbf{v}$ , where  $\times$  denotes a vector product, follows from the conservation of total momentum<sup>11</sup>. Thus, as  $t \rightarrow \infty$  the distribution function approaches  $f^*(\mathbf{r}, \mathbf{v}, t)$  where

$$f^*(\mathbf{r}, \mathbf{v}, t) = W_W(\mathbf{r})e^{[a(\mathbf{r}, t) + \mathbf{b}(\mathbf{r}, t) \cdot \mathbf{v} + c(\mathbf{r}, t)\mathbf{v}^2]}, \quad (74)$$

and for this distribution  $J(f^*, f^*) = 0$ . The condition that  $f^*$  be a solution of the Boltzmann equation also requires that

$$\frac{\partial f^*}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{r}} f^* + \frac{\mathbf{F}}{m} \cdot \nabla_{\mathbf{v}} f^* = 0 \quad (75)$$

inside the vessel containing the gas. If the container is at rest, and if the external force derives from a fixed potential  $\phi_{ext}(\mathbf{r})$ , it follows that at long times  $f^*$  must have the form<sup>12</sup>

$$f^*(\mathbf{r}, \mathbf{v}, t) = AW_W(\mathbf{r})e^{-\beta[\frac{1}{2}m\mathbf{v}^2 + \phi_{ext}(\mathbf{r})]}, \quad (76)$$

---

<sup>11</sup>Since  $\mathbf{r} \times \mathbf{v} + \mathbf{r} \times \mathbf{v}_1 = \mathbf{r} \times \mathbf{v}' + \mathbf{r} \times \mathbf{v}'_1$  follows from  $\mathbf{v} + \mathbf{v}_1 = \mathbf{v}' + \mathbf{v}'_1$ .

<sup>12</sup>Note that if the gas particles make specular collisions with the walls of the container, then the Maxwell-Boltzmann form is preserved by the interactions with the walls and there is no energy exchange between the walls and the gas, itself. The conclusions reached in the discussion to follow also apply to this case except that  $\beta$  is not determined by the walls, but rather by the total energy of the system.

where  $A$  and  $\beta$  are constants. We can determine  $A$  from the normalization of  $f^*$  as

$$A = nC(\beta) \left( \frac{\beta m}{2\pi} \right)^{\frac{d}{2}}, \quad (77)$$

where  $n = N/V$  is the number density of the gas, we have returned to a general number of spatial dimensions,  $d$ , and

$$C(\beta) = V \left[ \int d\mathbf{r} W_W(\mathbf{r}) e^{-\beta\phi_{ext}(\mathbf{r})} \right]^{-1}. \quad (78)$$

Finally, if the walls are maintained at temperature  $T_W$ , the thermostat condition requires that  $\beta = \beta_W = (k_B T_W)^{-1}$  when the distribution function reaches a stationary solution of the form Eq. (76).

We have therefore shown, by means of the  $H$ -Theorem that if the walls are maintained at a constant temperature  $T_W$ , and the particle-wall interaction satisfies the thermostat condition, *or* if there is no flux of energy at the walls, the distribution function for the gas approaches the Maxwell-Boltzmann equilibrium distribution function as  $t \rightarrow \infty$ .

## 2.4 The Objections of Zermelo and of Loschmidt

By means of the Boltzmann transport equation, and its corollary, the  $H$ -Theorem, Boltzmann appears to have found the connection between macroscopic thermodynamic behavior of dilute gases, and the microscopic properties of the particles composing the gas and their interactions with each other and with the boundaries. Boltzmann's solution addresses, in particular, the irreversible increase of entropy or the decrease of free energy in the approach of the gas to an equilibrium state. This is a very problematic result, however satisfactory it may be from the point of view of thermodynamics. In fact the irreversibility of the Boltzmann equation and the  $H$ -Theorem stand in contradiction to the fundamental reversibility of the laws of mechanics that govern all of the interactions between the various constituent parts of the system.

Let us consider the case where the gas particles make specular, elastic collisions with the walls, since this simplifies the argument without affecting the basic points. The laws of mechanics require that the form of the equations describing the motion of a system be invariant under the transformation

$t \rightarrow -t$ ;  $\mathbf{v} \rightarrow -\mathbf{v}$ ; and  $\mathbf{r} \rightarrow \mathbf{r}$ . Inspection of the Boltzmann equation for the behavior of the gas away from the boundary

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{r}} f + \frac{\mathbf{F}}{m} \cdot \nabla_{\mathbf{v}} f = J(f, f) \quad (79)$$

shows that each term on the left hand side of this equation changes in sign under this transformation, while the right hand side does not. Therefore, no matter what else might be correct about the Boltzmann equation, the statement that it is a direct consequence of the dynamics of the particles making up the gas and the vessel is not correct. This point was made sharply in the form of two interesting objections to the  $H$ -Theorem, one by Loschmidt in 1876, shortly after the publication of the  $H$ -theorem in 1872 by Boltzmann, the other in 1896 by Zermelo.

Loschmidt argued that the  $H$ -Theorem must be in violation of Newtonian mechanics because if the  $H$ -function were to decrease in time for a mechanical system, it would have to increase in time for the time reversed motion. Since any motion and its time reversed form are both equally valid solutions of the Newtonian equations, there is no way that the  $H$ -Theorem could be correct for both motions, unless of course the  $H$ -function were a time independent constant, which it is not. The present day understanding of this paradox is that it can be resolved by considering the statistical nature of the Boltzmann equation. Indeed, on a microscopic time scale the  $H$ -function is equally likely to be instantaneously increasing as decreasing, but if it is increasing it will do so for extremely short time only and on longer time scales it will be decreasing with overwhelming probability. One way of understanding this is by realizing that in an equilibrium ensemble states that are arbitrarily far removed from equilibrium do occur, be it with unimaginably small weights. However the decrease of weight of these states with distance from equilibrium (precisely this is quantified by the exponential of the  $H$ -function!) is so extremely rapid, that almost all trajectories corresponding to the nonequilibrium state under consideration will turn around into the direction of decreasing  $H$  virtually immediately. A nice illustration of this has been given by Ehrenfest in his treatment of the urn (or dog-flee) model[1]. Although this is a purely stochastic model it nicely catches the idea of almost all far-from-equilibrium states being reached at a point where the decrease of entropy is turning around. Finally, one has to assume that systems that are prepared by brute force in a far from equilibrium state will on average behave in the same way as systems that supposedly reach this

Table 1: The recurrence time,  $t_r$ , for one percent density fluctuations in a spherical volume of radius  $a$  in air at standard temperature and pressure.

$a$ [cm]	$t_r$ [sec]
$\sim 10^{-5}$	$\approx 10^{-11}$
$\sim 2.5 \times 10^{-5}$	$\approx 1$
$\sim 3 \times 10^{-5}$	$\approx 10^6$
$\sim 5 \times 10^{-5}$	$\approx 10^{68}$
$\sim 1$	$\approx 10^{10^{14}}$

state through an enormous fluctuation.

Zermelo's objection was somewhat more subtle. It relied upon a mathematical result of Poincaré called the *Recurrence Theorem*<sup>13</sup> Poincaré's theorem states that any isolated, finite mechanical system is *recurrent*. This means the following: Imagine a system of a finite number of particles contained in a finite container and isolated from the rest of the universe. Now imagine some initial state specifying the positions and momenta of all of the particles in the system. The recurrence theorem states that the mechanical system will eventually come arbitrarily close to its initial state, and this will happen over and over again given enough time. Thus even if the  $H$ -function were to decrease for some time in the motion of a system, it would eventually have to increase in order to come arbitrarily close to its previous higher value. Of course, the time needed for a macroscopic system to return sufficiently close to some initial state is ridiculously long, but that in no way invalidates the mathematical argument advanced by Zermelo. To get some feeling for the enormous times associated with the Poincaré recurrence theorem we present in Table 1 some results obtained by Chandrasekhar for the recurrence of a relatively mild fluctuation in density (which will occur extremely more frequently than an actual Poincaré recurrence!).

Given the evident problems reconciling the Boltzmann equation and its consequences with Newtonian mechanics, but yet the evident parallel between the equation and the laws of thermodynamics, we must now determine in what sense the Boltzmann equation can be understood as a correct

---

<sup>13</sup>This theorem was anticipated by Nietzsche's Doctrine of Eternal Return. Nietzsche's argument was not unlike that of Poincaré's proof but much less precise. See S. Brush.

description of the non-equilibrium behavior of a dilute gas. The answer, as we have indicated previously, is that it is a statement about the probable behavior of a gas, rather than about the individual behavior of any particular realization of the gas. To get some physical feeling for this situation imagine an isolated room containing a macroscopic number of molecules of a gas, and imagine an initial situation specified only by the condition that all of the molecules are confined to a small region of the room, say a corner. We would expect to see the gas molecules spread throughout the room, and this is indeed the typical behavior. However we can also imagine initial arrangement of the positions and velocities of the molecules such that for an interval of time, the gas becomes more compressed in the corner, or even moves without too much change into the opposite corner of the room. Any mechanically possible behavior could, in principle, occur, however unlikely or bizarre. But for such behavior to occur the requirements on the initial phases of the molecules are so special that the probability they will occur in a physical realization of the specified macroscopic initial state is effectively zero, and we never see such things. This example illustrates the statistical nature of the Boltzmann equation. It is a correct description of the probable non-equilibrium behavior of a gas, even though it is not a direct consequence of the laws of mechanics, and is, in some ways, in contradiction with these laws.

## 2.5 The Kac Ring Model. What is the Source of Irreversibility in the Boltzmann Equation?

As we stated in the derivation of the Boltzmann equation, a non-mechanical assumption is employed to obtain an expression for the rate of direct and of restituting collisions. This assumption was the *Stosszahlansatz*. It would be very illuminating if there were a model system that would be simple enough to illustrate the statistical nature of the *Stosszahlansatz* and to show how it might be understood as resulting from the application of statistical arguments for determining the probable behavior of an ensemble of identical reversible mechanical systems differing only in the initial phases of the constituent particles. Such a useful and clarifying model was invented by M. Kac and we turn to it now. It is known as the *Kac Ring Model*.

Consider a circle with a large but discrete number,  $N$ , of lattice sites placed along its circumference. Suppose further that there are  $M < N$

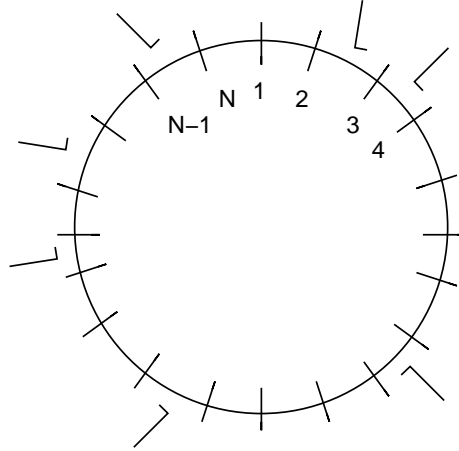


Figure 5: The Kac Ring Model. The sites on the ring are indicated and the markers in front of some of the sites are indicated by the check marks.

markers placed between adjacent sites, as illustrated in Fig. 5. The markers may be distributed along the circle in any way one might think of, the only rule being that there is at most one marker between any two adjacent sites. The markers are fixed once they are placed around the circle. Now, on each of the  $N$  lattice sites there is exactly one bead. The beads come in two colors, black and white, and the  $N$  black or white beads are distributed along the circle in some way. The dynamics of the model takes place at discrete time steps, and at each time step all the beads move exactly one step to the next lattice site, in a clockwise direction along the circumference of the circle. If the bead passes one of the markers in a step from one site to the next, it changes its color to the other one. That is all there is to the model and its dynamics. The dynamics are clearly reversible: If we start from some initial state and run the dynamics for  $t$  steps, we can recover the initial state by simply running the dynamics counter-clockwise for  $t$  steps. Furthermore, no matter what the initial arrangement of the beads might be, there is a Poincaré recurrence of  $2N$  steps, since at the  $2N$ -th step each bead will have passed every marker twice and be back to its initial position and color. Of course certain arrangements of the colors and markers may have shorter recurrence times as well, but  $2N$  is a universal recurrence time for this model.

The macroscopic quantity of interest will be taken to be the total number of beads of each color. Let  $B(t)$  and  $W(t)$  denote the number of black

beads and white beads, respectively, after  $t$  steps. We wish to determine an expression for  $\Delta(t) = B(t) - W(t)$ . We will do this in two ways: first, by using an argument employing the *Stosszahlansatz*; and second, by solving the exact equations of motion and then applying the methods of statistical mechanics to the exact dynamics. The following two steps are mechanically correct, in any case: The total number of beads is  $N$ ,

$$B(t) + W(t) = N. \quad (80)$$

If we denote by  $b(t)$  and  $w(t)$  the number of black beads and the number of white beads, respectively, at time  $t$ , about to change color in the next step, we find that

$$\begin{aligned} B(t+1) &= B(t) - b(t) + w(t), \\ W(t+1) &= W(t) - w(t) + b(t) \quad \text{or} \\ \Delta(t+1) &= \Delta(t) - 2(b(t) - w(t)). \end{aligned} \quad (81)$$

This is as far as we can go without examining the beads around the circle to learn which ones are about to change color. This is, of course, similar to our problem of trying to count the number of collisions taking place in time  $\delta t$  when we derived the Boltzmann equation. There we resorted to the *Stosszahlansatz* to get a number. Here we do the same: We say that the number of black or white beads about to change color is equal to the total number of beads of each color multiplied by the fraction of sites that have a marker in front of them, namely  $\mu = M/N$ . This is the analog of the *Stosszahlansatz* for the Ring Model, and it introduces the irreversibility into the equation for  $\Delta(t)$ . That is

$$\begin{aligned} b(t) &= \mu B(t), \quad w(t) = \mu W(t) \quad \text{and} \\ \Delta(t+1) &= (1 - 2\mu) \Delta(t). \end{aligned} \quad (82)$$

The solution of Eq. (82) for  $\Delta(t)$  is obviously

$$\Delta(t) = (1 - 2\mu)^t \Delta(0). \quad (83)$$

Equation (83) predicts a monotonic decay of  $|\Delta(t)|$  to its “equilibrium” value, zero. There is no sign in this equation of either the reversibility of the model or of the Poincaré recurrence phenomena. How then can it possibly be correct, in any sense whatever?

To answer this question we now turn to a more exact description of the dynamics of the model. We define some additional quantities,  $\eta_i(t)$  and  $\epsilon_j$ , where  $\eta_i(t) = 1$  if the bead at site  $i$  at time  $t$  is black, and  $\eta_i(t) = -1$  if the bead at site  $i$  at time  $t$  is white. Also  $\epsilon_i = 1$  if there is no marker in front of site  $i$  and  $\epsilon_i = -1$  if there is a marker in front of site  $i$ . Using these two definitions we can easily see that the microscopic equation of motion for this model is simply

$$\eta_i(t+1) = \epsilon_{i-1}\eta_{i-1}(t). \quad (84)$$

This simple equation embodies all of the dynamics of the model. By iteration we obtain

$$\eta_{i+1}(t+1) = \epsilon_i\epsilon_{i-1}\cdots\epsilon_{i-t}\eta_{i-t}(0). \quad (85)$$

By summing  $\eta_i(t)$  over all the sites we obtain an expression for  $\Delta(t)$ , as

$$\Delta(t+1) = \sum_i \epsilon_i\epsilon_{i-1}\cdots\epsilon_{i-t}\eta_{i-t}(0). \quad (86)$$

Notice that the Poincaré recurrence time  $2N$  appears as a consequence of Eqs. (85) or (86). That is  $\eta_i(2N) = \eta_i(0)$  and  $\Delta(2N) = \Delta(0)$  since for this case every  $\epsilon_j$  appears twice in each product of  $\epsilon$ 's.

Now consider an ensemble of Kac Rings, each with the same distribution of beads around the circle, but differing in the placement of the markers. In fact we shall prepare the ensemble in such a way that the fraction of members of the ensemble with a marker in front of any site has the value  $\mu$  as defined above. This means that the average number of markers is still  $M$  but individual members of the ensemble may have more or less markers than  $M$  in accordance with the above prescription. Under these circumstances the  $\epsilon_i$  may be considered to be identical individually distributed random variables with average values  $\langle \epsilon_i \rangle = (1 - \mu)(+1) + \mu(-1) = (1 - 2\mu)$ . Using this result we may compute all of the averages, being careful to note if a particular  $\epsilon_j$  appears more than once in the product. For this reason, we obtain different results if  $0 < t \leq N$  and if  $N < t \leq 2N$ . For  $t$  in the interval  $0 < t \leq N$  we find

$$\begin{aligned} \langle \Delta(t) \rangle &= \sum_i \langle \epsilon_i\epsilon_{i-1}\cdots\epsilon_{i-t} \rangle \Delta(0), \\ &= (\langle \epsilon_i \rangle)^t \Delta(0) = (1 - 2\mu)^t \Delta(0). \end{aligned} \quad (87)$$

this is exactly the result that we obtained using the *Stosszahlansatz*, but its validity is now restricted to a fixed range of time  $t$ . On the other hand for

$N < t \leq 2N$ , the counting is somewhat different. We have to ignore all markers that appear twice in the products since they give a factor of unity. Thus, suppose  $t = N + s$  then any markers between site  $N$  and site  $s - 1$  would be counted twice. In that case there will be  $N - s = 2N - t$  terms in the products of the  $\epsilon$ 's, and, as a result

$$\langle \Delta(t) \rangle = (1 - 2\mu)^{2N-t} \Delta(0), \quad (88)$$

a result quite consistent with the Poincaré recurrence theorem, since the initial value of  $\Delta$  is always recovered after exactly  $2N$  steps. The exponential decay of  $\langle \Delta(t) \rangle$ , as described by Eq. (87) is the average result for our ensemble for times less than the number of sites. Of course if  $N$  is a huge number, say on the order of  $10^{23}$ . it would seem unlikely that we would see anything else.

To make this point more precise we need to ask about the size of the fluctuations about the results just obtained. If  $N$  is very large, we would expect these results to be *typical* in the sense that the fluctuations about them are very small. This has been discussed by Kac, who has shown that the fluctuations are indeed very small for large  $N$ . However, there are certainly members of the ensemble where there are large regions free of, or filled with, markers. For these members we would expect to see non-Boltzmann behaviors.

## 2.6 The Chapman-Enskog Solution

In this subsection we will discuss solutions of the Boltzmann equation that are relevant for our understanding of transport phenomena in dilute gases. We especially discuss the so-called normal or Chapman-Enskog solution of the equation that is appropriate for a gas that is close to a local equilibrium state at every point in space. That is to say, the state of the gas is such that at every point one can define a local density, mean velocity, and temperature, but these quantities may vary on some large scale over the container. The Chapman-Enskog solution to the Boltzmann equation is predicated upon the existence of such a situation and then proceeds to derive, by methods to be discussed here, the macroscopic, Navier-Stokes, equations of fluid dynamics together with expressions for the transport coefficients appearing in them in terms of the dynamics of binary collisions taking place in a dilute gas.

Consider the Boltzmann equation,

$$\frac{\partial f(\mathbf{r}, \mathbf{v}, t)}{\partial t} + \nabla_{\mathbf{r}} \cdot (\mathbf{v}f(\mathbf{r}, \mathbf{v}, t)) + \nabla_{\mathbf{v}} \cdot (\mathbf{F}f(\mathbf{r}, \mathbf{v}, t)) = J(f, f), \quad (89)$$

where we introduced the symbol  $\mathbf{F}$  for the external force per unit mass. We are interested in situations where both the time derivative and the gradients on the left hand side are very small. This implies the right hand side must be very small as well, suggesting some kind of expansion around solutions for which it vanishes. As argued in the first section (see Eq. (73)), these solutions are the *local equilibrium distributions*

$$\tilde{f}_{loc}(\mathbf{r}, \mathbf{c}, t) = n(\mathbf{r}, t) \left( \frac{\beta(\mathbf{r}, t)m}{2\pi} \right)^{d/2} e^{-\beta(\mathbf{r}, t)mc^2(\mathbf{r}, t)/2}. \quad (90)$$

Here  $\beta(\mathbf{r}, t) = (k_B T(\mathbf{r}, t))^{-1}$  and  $\mathbf{c}(\mathbf{r}, t) = \mathbf{v} - \mathbf{u}(\mathbf{r}, t)$ , with  $\mathbf{u}(\mathbf{r}, t)$  the local average velocity.

The essence of the Chapman-Enskog solution is: (1) to suppose that after a few mean free times the collisions taking place in the gas produce a state that is fully determined by the hydrodynamic densities  $n(\mathbf{r}, t)$ ,  $\mathbf{u}(\mathbf{r}, t)$  and  $T(\mathbf{r}, t)$  (apart from the boundary conditions), and then, (2) to express the solution of the Boltzmann equation as an expansion in powers of the small parameter  $\mu = \ell/L$ , with  $\ell$  the mean free path between collisions and  $L$  a characteristic macroscopic length, typically a distance over which gradients of the hydrodynamic densities induce notable changes. Acting upon distribution functions that are not close to local equilibrium ones the left hand side of Eq. (90) under the required conditions is of order  $\mu$  (and possibly containing higher order terms) compared to the right hand side. In accordance with this approach we write

$$f(\mathbf{r}, \mathbf{v}, t) = \tilde{f}_{loc}(\mathbf{r}, \mathbf{c}, t) \left[ 1 + \mu \tilde{\Phi}_1(\mathbf{r}, \mathbf{v}, t) + \mu^2 \tilde{\Phi}_2(\mathbf{r}, \mathbf{v}, t) + \dots \right], \quad (91)$$

Here the functions  $\tilde{\Phi}_n(\mathbf{r}, \mathbf{v}, t) \equiv \mu^n \tilde{\Phi}_n(\mathbf{r}, \mathbf{v}, t)$  are to be determined from the Boltzmann equation, together with, and as functions of the hydrodynamic variables,  $n(\mathbf{r}, t)$ ,  $T(\mathbf{r}, t)$  and  $\mathbf{u}(\mathbf{r}, t)$  and powers of gradients acting upon these.

The most straightforward way of solving the Boltzmann equation perturbatively now would be expanding both the distribution function and the

hydrodynamic fields into series in powers of  $\mu$ , determining the zeroth order from the equation

$$J(f_0, f_0) = 0, \tag{92}$$

with given initial values for the hydrodynamic fields, and solving successively for subsequent orders  $\Phi_n$ , together with subsequent contributions to the time dependent hydrodynamic fields. This procedure is known as the *Hilbert expansion*. It works in principle, but to reproduce e.g. a solution of the nonlinear Navier-Stokes equations over a not even very long time one already has to iterate over many orders <sup>14</sup>. In contrast, the Chapman-Enskog solution reproduces the full Navier Stokes equations already after two iterations. The price paid for this is that it cannot be considered a fully systematic expansion in powers of  $\mu$ , yet, for each order in  $\mu$  the procedure is uniquely defined. To be more specific: in the  $n$ -th order Chapman-Enskog solution the starting point likewise is Eq. (92) and, like in the Hilbert solution the functions  $\Phi_1 \cdots \Phi_n$  are solved successively. The difference however, is that the time and position dependence of the hydrodynamic fields in terms of which all these functions are expressed are determined from the full set of equations through order  $n$ . Each of the functions  $\Phi_j$  is expressed in terms of local density, velocity and temperature, including up to  $n$ -th order gradients of these, but the actual time and position dependence of these fields can be determined only after deriving the explicit form of the hydrodynamic equations through order  $n$ .

As a consequence of this procedure the time derivative of the distribution

---

<sup>14</sup>This situation is addressed, mathematically, by *Hilbert's Uniqueness Theorem*. The normal solution method which we develop in this section leads to a solution of the Boltzmann equation that depends *only* upon the initial values of the hydrodynamic fields,  $n(\mathbf{r}, 0), \mathbf{u}(\mathbf{r}, 0), T(\mathbf{r}, 0)$ . Hilbert's Uniqueness Theorem states that if the distribution function  $f(\mathbf{r}, \mathbf{v}, t)$  can be expanded in a convergent series in powers of some small parameter, then this function is *uniquely* determined by the initial values of the hydrodynamic fields. Under such circumstances, we would obtain the same distribution function whether we were to expand the distribution function about a total equilibrium distribution in powers of some small parameter characterizing the deviations from total equilibrium, or about local equilibrium in powers of  $\mu$ . This all assumes that the series expansions converge. In many cases the expansions are likely to be asymptotic rather than convergent, Hilbert's theorem is then of little use, the mathematical foundations of the method are not secure, and one must rely on physical intuition and experiment to validate the procedure. However there are examples, notably in the case of hard-sphere gases or other potential models with finite total cross sections, and some Lorentz models, where one can prove the convergence of some of the expansions.

function, appearing on the left hand side of (89) is not simply of order  $\mu$ . Through the time derivatives of the hydrodynamic fields terms of orders  $\mu$  through  $\mu^n$  are generated, even though in the expansion procedure the fields themselves are treated at all times as though being of order  $\mu^0$ , while each spatial derivative is treated as bringing forth an order  $\mu$ .

The local equilibrium distribution, which is the general solution of (92), contains the full hydrodynamic fields for all times. Chapman and Enskog found out the solution method can be made unique in a simple and useful way by requiring that this zeroth order solution alone reproduces the full hydrodynamic densities for all times and positions, viz.

$$\int \psi(\mathbf{v})f(\mathbf{r}, \mathbf{v}, t)d\mathbf{v} = \int \psi(\mathbf{v})\tilde{f}_{loc}(\mathbf{r}, \mathbf{c}, t)d\mathbf{v}, \quad (93)$$

where the  $\psi$  is any linear combination of the conserved quantities,  $m, \mathbf{v}$  and  $mv^2/2$ , identified just below Eq. (73). This condition implies that

$$\int \psi(\mathbf{v})\Phi_j\tilde{f}_{loc}(\mathbf{r}, \mathbf{c}, t)d\mathbf{v} = 0, \quad \text{for } j = 1, 2, \dots \quad (94)$$

### 2.6.1 The $\mu$ -expansion of the Boltzmann equation

If we insert the proposed solution, Eq. (91) into Eq. (89), we obtain

$$\left[ \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla + \mathbf{F} \cdot \nabla_{\mathbf{v}} \right] \tilde{f}_{loc} \left( 1 + \mu\tilde{\Phi}_1 + \dots \right) = J(\tilde{f}_{loc}, \tilde{f}_{loc}) + \mu\tilde{f}_{loc}L(\mathbf{c})\tilde{\Phi}_1 + \mu^2 \left[ \tilde{f}_{loc}L(\mathbf{c})\tilde{\Phi}_2 + J(\tilde{f}_{loc}\tilde{\Phi}_1(\mathbf{v}), \tilde{f}_{loc}\tilde{\Phi}_1(\mathbf{v}_1)) \right] + O(\mu^3) \quad (95)$$

where

$$\tilde{f}_{loc}L(\mathbf{c})\tilde{\Phi}_1 = J(\tilde{f}_{loc}\tilde{\Phi}_1, \tilde{f}_{loc}) + J(\tilde{f}_{loc}, \tilde{f}_{loc}\tilde{\Phi}_1), \quad (96)$$

and we used  $\nabla$  for  $\nabla_{\mathbf{r}}$  as we will keep doing in the sequel. The order in powers of the parameter  $\mu$  of the terms appearing on the right hand side of Eq. (95) is immediately clear and we can also easily assess the ordering of terms on the left hand side. As noted in the previous subsection both gradient terms raise the order of terms they act upon by one power of  $\mu$ . The time derivatives of the various contributions to  $\tilde{f}(\mathbf{r}, \mathbf{v}, t)$  all result from the time dependence of the hydrodynamic densities. These in general contain higher order gradients and powers of gradients, so it is reasonable to suppose

that the time derivatives of these densities,  $\rho_i(\mathbf{r}, t)$  are given by expansions in powers of  $\mu$ , which we write in the form

$$\frac{\partial \rho_i(\mathbf{r}, t)}{\partial t} = \left( \frac{\partial \rho_i(\mathbf{r}, t)}{\partial t} \right)_1 + \left( \frac{\partial \rho_i(\mathbf{r}, t)}{\partial t} \right)_2 + \dots, \quad (97)$$

where  $\left( \frac{\partial}{\partial t} \right)_j$  should be understood to represent terms of order  $\mu^j$ .

### 2.6.2 The distribution function to zeroth order in $\mu$ and the Euler equations

The left hand side of the Boltzmann equation, as we argued above, is at least of order  $\mu$ . Thus to zeroth order in  $\mu$  the Boltzmann equation, as noted before is simply

$$J(f_0, f_0) = 0 \quad (98)$$

where  $f_0$  is the zeroth order approximation to the distribution function. We have already identified the solution for this with the local equilibrium distribution function

$$f_0(\mathbf{r}, \mathbf{v}, t) \equiv \tilde{f}_{loc}(\mathbf{r}, \mathbf{c}, t), \quad (99)$$

where the local equilibrium distribution has the Maxwell-Boltzmann form and is given by Eq. (90).

To first order in  $\mu$  we obtain

$$\begin{aligned} & \left[ \frac{\partial \tilde{f}_{loc}(\mathbf{r}, \mathbf{c}, t)}{\partial t} \right]_1 + \mathbf{v} \cdot \nabla_{\mathbf{r}} \tilde{f}_{loc}(\mathbf{r}, \mathbf{c}, t) + \mathbf{F} \cdot \nabla_{\mathbf{v}} \tilde{f}_{loc}(\mathbf{r}, \mathbf{c}, t) = \\ & = \int d\hat{\boldsymbol{\sigma}} \int d\mathbf{v}_1 |\mathbf{g} \cdot \hat{\boldsymbol{\sigma}}| \tilde{f}_{loc}(\mathbf{r}, \mathbf{c}, t) \tilde{f}_{loc}(\mathbf{r}, \mathbf{v}_1, t) [\Phi_1(\mathbf{r}, \mathbf{v}', t) + \Phi_1(\mathbf{r}, \mathbf{v}_1', t) \\ & - \Phi_1(\mathbf{r}, \mathbf{v}, t) - \Phi_1(\mathbf{r}, \mathbf{v}_1, t)]. \end{aligned} \quad (100)$$

The subscript on the time derivative in the above equation serves as a reminder that we retain only terms of order  $\mu$  when calculating the time derivative.

By taking the inner products of this with 1,  $m\mathbf{v}$  and  $mv^2/2$  we obtain

the  $d + 2$  Euler, or ideal fluid equations of flow,

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot (\rho \mathbf{u}) \quad (101)$$

$$\rho_m \left( \frac{\partial \mathbf{u}}{\partial t} \right)_1 + \mathbf{u} \cdot \nabla \mathbf{u} = -\nabla p + \rho_m \mathbf{F}; \quad (102)$$

$$\frac{d}{2} n k_B \frac{dT}{dt} + n k_B T (\nabla \cdot \mathbf{u}) = 0, \quad (103)$$

with the local equilibrium pressure  $p$  satisfying the ideal gas law

$$p(\mathbf{r}, t) = n(\mathbf{r}, t) k_B T(\mathbf{r}, t) \quad (104)$$

These equations reproduce the first order time derivatives of the hydrodynamic densities. They describe the motion of an ideal low density fluid, which is both *inviscid* and *adiabatic*. In their derivation we used the vanishing of the inner product of the collision invariants  $1, \mathbf{v}$  and  $v^2$  with the collision operator on the right hand side and we employed the continuity equation (101) to simplify the other two equations somewhat.

### 2.6.3 The distribution function to order $\mu$ : The Navier-Stokes equations

Now we turn our attention to obtaining the first order correction to the local equilibrium function,  $\Phi_1$  which will lead us to the Navier-Stokes equations of fluid dynamics, and provide us with expressions for the coefficients of viscosity and thermal conductivity in terms of the intermolecular forces that govern the binary collisions taking place in the fluid. Starting point is Eq. (100). To eliminate the time derivatives from this, to order  $\mu$  we first notice that the local equilibrium distribution depends on time only through the hydrodynamic densities and then use the Euler equations derived in the previous subsection to reexpress the time derivatives of these in terms of gradients.

Inserting the resulting expressions in the left hand side of Eq. (100) we obtain the following integral equation for  $\Phi_1$

$$\begin{aligned} -\tilde{f}_{loc}(\mathbf{r}, \mathbf{c}, t) \left[ \beta(\mathbf{r}, t) m \left( \mathbf{c}\mathbf{c} - \frac{c^2}{d} \mathbf{1} \right) : \nabla \mathbf{u} + \left( \frac{\beta m}{2} c^2 - \frac{d+2}{2} \right) \mathbf{c} \cdot \nabla \ln T(\mathbf{r}, t) \right] \\ = \tilde{f}_{loc}(\mathbf{r}, \mathbf{c}, t) L(\mathbf{r}, \mathbf{c}, t) \Phi_1. \end{aligned} \quad (105)$$

The linearized Boltzmann collision operator,  $L(\mathbf{r}, \mathbf{c}, t)$ , appearing in this equation was defined in Eq. (96). This operator is symmetric and negative

definite outside of the space of conserved quantities. It depends implicitly on the temperature, density and average velocity at the position  $\mathbf{r}$ . Eigenfunctions belonging to different eigenvalues are orthogonal to each other, under an inner product defined through

$$(f, g)_{rt} = \int d\mathbf{c} \frac{\tilde{f}_{loc}(\mathbf{r}, \mathbf{c}, t)}{n(\mathbf{r}, t)} f^*(\mathbf{c}) g(\mathbf{c}). \quad (106)$$

The proof of the negative definite property follows the lines of the proof of the  $H$ -Theorem, and is simply the statement that

$$\begin{aligned} & \int d\mathbf{c} \tilde{f}_{loc}(\mathbf{r}, \mathbf{c}, t)(\mathbf{c}) \psi^*(\mathbf{c}) L(\mathbf{r}, \mathbf{c}, t) \psi(\mathbf{c}) = \\ & - \frac{1}{4} \int d\mathbf{c} \int d\mathbf{c}_1 \int d\hat{\boldsymbol{\sigma}} B(\mathbf{g}, \hat{\boldsymbol{\sigma}}) \tilde{f}_{loc}(\mathbf{c}) \tilde{f}_{loc}(\mathbf{c}_1) \times \\ & \|\psi(\mathbf{c}) + \psi(\mathbf{c}_1) - \psi(\mathbf{c}') - \psi(\mathbf{c}'_1)\|^2 < 0, \end{aligned} \quad (107)$$

provided  $\psi(\mathbf{c})$  is not a linear combination of quantities conserved in a binary collision.

Eq. (105) can be solved as

$$\Phi_1(\mathbf{r}, \mathbf{c}, t) = -L^{-1}(\mathbf{r}, \mathbf{c}, t) \left[ \beta(\mathbf{r}, t) m \left( \mathbf{c} \mathbf{c} - \frac{c^2}{d} \mathbf{1} \right) : \nabla \mathbf{u} + \left( \frac{\beta m}{2} c^2 - \frac{d+2}{2} \right) \mathbf{c} \cdot \nabla \ln T(\mathbf{r}, t) \right], \quad (108)$$

with  $L^{-1}$  defined on the subspace of functions that are orthogonal to 1,  $\mathbf{c}$  and  $c^2$ . The Boltzmann equation may now be expanded through second order in  $\mu$  as

$$\begin{aligned} & \left[ \frac{\partial \tilde{f}_{loc}(\mathbf{r}, \mathbf{c}, t)}{\partial t} \right]_1 + \left[ \frac{\partial \tilde{f}_{loc}(\mathbf{r}, \mathbf{c}, t)}{\partial t} \right]_2 + \mathbf{v} \cdot \nabla \tilde{f}_{loc}(\mathbf{r}, \mathbf{c}, t) (1 + \Phi_1(\mathbf{r}, \mathbf{c}, t)) \\ & + \mathbf{F} \cdot \nabla_{\mathbf{v}} \tilde{f}_{loc}(\mathbf{r}, \mathbf{c}, t) (1 + \Phi_1(\mathbf{r}, \mathbf{c}, t)) \\ & = J(f_{loc}(1 + \Phi_1), f_{loc}(1 + \Phi_1)) + J(f_{loc} \Phi_2, f_{loc}) + J(f_{loc}, f_{loc} \Phi_2) \end{aligned} \quad (109)$$

By taking the inner product of this equation with  $m\mathbf{v}$  respectively  $\frac{1}{2}mv^2$  one obtains the Navier-Stokes equations,

$$\rho_m \frac{d\mathbf{u}}{dt} = -\nabla p + \mathbf{F} \rho_m + \nabla \cdot [2\eta \mathbf{D} - \frac{1}{d} \mathbf{1} (\nabla \cdot \mathbf{u})], \quad (110)$$

$$\frac{d}{2} n k_B \frac{dT}{dt} + n k_B T (\nabla \cdot \mathbf{u}) = \nabla \cdot \lambda \nabla T + 2\eta [\mathbf{D} : \mathbf{D} - \frac{1}{d} (\nabla \cdot \mathbf{u})^2], \quad (111)$$

with  $d/dt \equiv \partial/\partial t + \mathbf{u} \cdot \nabla$  the comoving time derivative and  $\mathbf{D}$  the symmetrized gradient of the velocity field, most easily defined through its components,

$$D_{ij} = \frac{1}{2} \left[ \frac{\partial \mathbf{u}_i}{\partial x_j} + \frac{\partial \mathbf{u}_j}{\partial x_i} \right]. \quad (112)$$

Furthermore the shear viscosity  $\eta$  and the heat conduction coefficient  $\lambda$  are given by

$$\eta(\mathbf{r}, t) = -\beta m (\mathbf{c}_x \mathbf{c}_y, L^{-1} \mathbf{c}_x \mathbf{c}_y)_{\mathbf{r}t} \quad (113)$$

$$\lambda(\mathbf{r}, t) = -k_B \left( \frac{[\beta m c^2 - d - 2] c_x}{2}, L^{-1} \frac{[\beta m c^2 - d - 2] c_x}{2} \right)_{\mathbf{r}t}. \quad (114)$$

Notice that in these equations there appears no bulk viscosity. This is a consequence of the fact that the microscopic pressure tensor for a dilute gas is purely kinetic. As a result the trace of this tensor is proportional to the kinetic energy, which is one of the conserved variables.

The evaluation of the transport coefficients for given interparticle potentials is a fairly complicated mathematical problem; for most interparticle potentials the inversion of the operator  $L$  is not straightforward and requires expansions in sets of orthogonal polynomials. There is an extensive literature on this[2, 3, 4] to which I refer the interested reader.

Extensions to mixtures can be made straightforwardly. The resulting set of transport coefficients can be shown to satisfy Onsager's reciprocal relations relating various pairs of transport coefficients to each other.

### 3 The Liouville and BBGKY Hierarchy Equations

Starting point for formal, and also for many heuristic derivations of kinetic equations usually are the BBGKY hierarchy equations. These are obtained from the Liouville equation,

$$\frac{\partial \rho(\Gamma, t)}{\partial t} + \mathcal{L}(\Gamma) \rho(t) = 0. \quad (115)$$

The Liouville operator can be decomposed into a kinetic part,  $\mathcal{L}_0(\Gamma)$  and a potential part  $\mathcal{L}_I(\Gamma)$ ,

$$\mathcal{L}(\Gamma) = \mathcal{L}_0(\Gamma) + \mathcal{L}_I(\Gamma). \quad (116)$$

which are given, respectively, by

$$\begin{aligned}\mathcal{L}_0(\Gamma) &= \sum_{i=1}^N \dot{\mathbf{r}}_i \cdot \frac{\partial}{\partial \mathbf{r}_i} = \sum_{i=1}^N \mathbf{v}_i \cdot \frac{\partial}{\partial \mathbf{r}_i}, \quad \text{and} \\ \mathcal{L}_I(\Gamma) &= \sum_{i=1}^N \dot{\mathbf{p}}_i^{int} \cdot \frac{\partial}{\partial \mathbf{p}_i} = - \sum_{i < j} \Theta_{i,j},\end{aligned}\tag{117}$$

where the latter operators are defined as<sup>15</sup>

$$\Theta_{i,j} = \frac{\partial \phi(|\mathbf{r}_i - \mathbf{r}_j|)}{\partial \mathbf{r}_i} \cdot \left[ \frac{\partial}{\partial \mathbf{p}_i} - \frac{\partial}{\partial \mathbf{p}_j} \right],\tag{118}$$

In the expression for  $\mathcal{L}_I$  only that part of the force on each particle  $i$  is included that is due to the interaction potential between this particle and all the other particles in the system. Interactions with the walls and other external forces, if present will give rise to additional terms in  $\mathcal{L}_0(\Gamma)$ , which either are of the form  $\sum_i \mathbf{F}_{ext}(\mathbf{r}_i) \cdot \partial / \partial \mathbf{p}_i$  or can be expressed as sums of wall collision operators. We have also assumed that there are only central forces between the particles. Equation (115) has the formal solution

$$\rho(\Gamma, t) = \mathcal{S}_{-t}(\Gamma) \rho(\Gamma, 0) = \rho(\Gamma(-t), 0),\tag{119}$$

where the time displacement operator  $\mathcal{S}_{-t}(\Gamma)$  is defined, formally, by

$$\mathcal{S}_{-t}(\Gamma) \equiv e^{-t\mathcal{L}(\Gamma)}.\tag{120}$$

The second identity in Eq. (119) is an immediate consequence of Liouville's theorem.

In practice the  $N$ -particle distribution function is hard to use and contains much more detailed information than one is interested to know. Often reduced  $n$ -particle distribution functions of the form

$$F_n(x_1, \dots, x_n, t) = \frac{N!}{(N-n)!} \int dx_{n+1} \dots dx_N \rho(x_1 \dots x_n, x_{n+1} \dots x_N, t),\tag{121}$$

---

<sup>15</sup>This form again presupposes pairwise additive potentials. In the case of  $p$ -body interactions  $\dot{\mathbf{p}}_i$  must be expressed likewise in terms of gradients of all interaction terms involving particle  $i$

with  $x_i = (\mathbf{r}_i, \mathbf{p}_i)$ , contain all the information one is really interested in. E.g. mass density and momentum density are contained in the one-particle distribution function, and for a system with pair interactions only, the pair distribution suffices for the energy density (it is needed for the potential energy density. For the kinetic energy density the one-particle distribution suffices, as we saw already in the previous section). By integrating all terms in the Liouville equation (115) over the coordinates of all particles except the first  $n$  ones and the arguments of  $\Theta_{i,j}$ , in case either of them is among  $1 \cdots n$  one obtains time evolution equations for the reduced distribution functions in the form

$$\left[ +\frac{\partial}{\partial t} + \mathcal{L}(\S_\infty, \cdots \S_n) \right] F_n(x_1, \dots, x_n, t) = \sum_{i=1}^n \int dx_{n+1} \Theta_{i,n+1} F_{n+1}(x_1, \dots, x_n, x_{n+1}, t). \quad (122)$$

These form a set of equations for  $n = 1, 2, \dots, N$ . The case where  $n = N$  is, of course, the Liouville equation since we must set  $F_{N+1} = 0$  in an  $N$ -particle system. The most important thing to notice here is that the right hand side of the equation for  $F_n$ , the reduced distribution function for  $n$  particles, contains the reduced distribution function  $F_{n+1}$ , namely the reduced function for  $n+1$  particles. Thus only the Liouville equation is closed, while the equations for reduced functions with fewer than  $N$  particles are not. This circumstance gives this set of equations its hierarchical structure. To use the BBGKY hierarchy equations one must find some way to close the hierarchy, that is, to express, for some value of  $n$  the function  $F_{n+1}$  in terms of lower order functions.

### 3.1 Low densities. The Boltzmann closure

For low-density systems apparently Boltzmann's Stosszahlansatz provides a way to close the hierarchy by expressing the pair distribution function for pre-collisional configurations as a product of two single particle distribution functions. This, however is less straightforward than it sounds. The first hierarchy equation is of the form

$$\begin{aligned} & \frac{\partial}{\partial t} f(\mathbf{r}_1, \mathbf{v}_1, t) + \nabla_{\mathbf{r}_1} \cdot (\mathbf{v}_1 f(\mathbf{r}_1, \mathbf{v}_1, t)) + \nabla_{\mathbf{v}_1} \cdot (\mathbf{F}(\mathbf{r}_1, \mathbf{v}_1) f(\mathbf{r}_1, \mathbf{v}_1, t)) = \\ & = \int d\mathbf{r}_2 \int d\mathbf{v}_2 \Theta(1, 2) F_2(\mathbf{r}_1, \mathbf{v}_1, \mathbf{r}_2, \mathbf{v}_2, t), \end{aligned} \quad (123)$$

where we identified  $F_1$  and  $f$ . Replacement of  $F_2$  by a product of two  $f$ 's does not give the Boltzmann equation, but the non-linear Vlasov equation. This is an interesting kinetic equation in its own right, which works well for weakly interacting particles. It is very important in plasma physics. However, it is a reversible equation and it does not satisfy an H-theorem. The problem with this factorization is that the  $\Theta$  operator acts on the pair distribution function for precisely such configurations where particles 1 and 2 are involved in an ongoing collision. In such situations correlations between their coordinates have built up, even though they may have been absent before the collision started.

In order to avoid this problem one should extrapolate the coordinates of 1 and 2 back to the instant just before their collision started and apply the Stosszahlansatz for these. A formal way to do this is by combining the first two hierarchy equations. The second one for this purpose only is needed over times in the order of the duration of a collision, therefore the term on the right hand side may be ignored and this equation may be written as

$$\frac{\partial F_2(1, 2, t)}{\partial t} = (\Theta_{1,2} - \mathcal{L}_0(1, 2)) F_2(1, 2, t). \quad (124)$$

This can be transformed into the integral equation

$$F_2(1, 2, t) = \mathcal{S}_{-t}^{(0)} F_2(1, 2, 0) + \int_0^t d\tau \mathcal{S}_{t-\tau}^{(0)} \Theta_{1,2} F_2(1, 2, \tau), \quad (125)$$

with  $\mathcal{S}_{-t}^{(0)} = \exp(-\mathcal{L}_0(1, 2)t)$ . By iterating this relation inside the integral one obtains

$$F_2(1, 2, t) = \mathcal{S}_{-t}^{(0)} F_2(1, 2, 0) + \int_0^t d\tau \int_0^\tau d\tau' \mathcal{S}_{t-\tau}^{(0)} \bar{T}_{1,2}(\tau') \mathcal{S}_{\tau'-\tau}^{(0)} F_2(1, 2, 0), \quad (126)$$

with the binary collision operator  $\bar{T}_{1,2}$  defined as

$$\bar{T}_{1,2}(t) = \Theta_{1,2} \delta(t) + \Theta_{1,2} \mathcal{S}_{-t}^{(0)} \Theta_{1,2} + \int_0^t d\tau \Theta_{1,2} \mathcal{S}_{-\tau}^{(0)} \Theta_{1,2} \mathcal{S}_{-t+\tau}^{(0)} \Theta_{1,2} + \dots \quad (127)$$

Since collisions last very shortly on the characteristic time scale of the mean duration between collisions, for most situations where the collision operator does not vanish the initial time  $-t$  will be before the start and the final time

$t = 0$  will be after the completion of the collision, To a good approximation then Eq. (126) may be replaced by

$$F_2(1, 2, t) = \mathcal{S}_{-t}^{(0)} F_2(1, 2, 0) + \int_0^t d\tau \mathcal{S}_{t-\tau}^{(0)} \bar{\mathbf{T}}_{1,2} \mathcal{S}_{-\tau}^{(0)} F_2(1, 2, 0), \quad (128)$$

with

$$\bar{\mathbf{T}}_{1,2} = \int_0^\infty dt \bar{\mathbf{T}}_{1,2}(t). \quad (129)$$

Now, inspecting this equation, one sees that the operator  $\bar{\mathbf{T}}_{1,2}$  on the one hand has to generate a term that performs the actual collision dynamics on the pair distribution function and on the other hand it has to produce a term that subtracts the free streaming term on the right hand side of Eq. (128). If one applies the Stosszahlansatz on  $F_2(1, 2, 0)$ , ignoring changes in  $F_1$  over distances in the order of a molecular diameter and over times of the order of the collision time one finds that the action of  $\bar{\mathbf{T}}_{1,2}$  reduces exactly to that of the Boltzmann collision operator. Finally, substituting the identity  $\Theta_{1,2} F_2(1, 2, -t) = \bar{\mathbf{T}}_{1,2} \mathcal{S}_{-t}^{(0)} F_2(1, 2, 0)$ , with  $t$  of the order of the duration of a collision, into the first hierarchy equation (123) one recovers the Boltzmann equation.

### 3.2 Hard spheres. Enskog and revised Enskog closure

A much studied model in kinetic theory is that of elastically colliding hard spheres. This has the great advantage that all collisions are strictly instantaneous and never involve more than two particles, even at very high densities. A further advantage is that, apart from contributions due to external potentials all energy is kinetic and therefore the single particle distribution function suffices for determining the hydrodynamic variables, also for all densities. In addition it has turned out that the transport coefficients of hard sphere systems, if properly interpreted give fairly accurate predictions for those of real systems, up to reasonably high densities[5].

The Liouville equation in the form (115) cannot be applied to hard sphere systems, as it requires a continuous and almost everywhere differentiable potential. However, thanks to the instantaneity of the collisions it can be replaced by a pseudo-Liouville equation[6],

$$\frac{\partial \rho(\Gamma, t)}{\partial t} + \mathcal{L}_0(\Gamma) \rho(\Gamma, t) = \sum_{i < j}^N \bar{\mathbf{T}}_-^{(hs)}(i, j) \rho(\Gamma, t). \quad (130)$$

Here the operators  $\bar{\mathbf{T}}_-^{(hs)}(i, j)$  are hard-sphere binary collision operators, defined as

$$\bar{\mathbf{T}}_-^{(hs)}(1, 2) = a^{d-1} \int_{\mathbf{v}_{12} \cdot \hat{\boldsymbol{\sigma}} > 0} d\hat{\boldsymbol{\sigma}} |\mathbf{v}_{12} \cdot \hat{\boldsymbol{\sigma}}| [\delta(\mathbf{r}_{12} - a\hat{\boldsymbol{\sigma}}) \mathbf{b}_{\hat{\boldsymbol{\sigma}}}(1, 2) - \delta(\mathbf{r}_{12} + a\hat{\boldsymbol{\sigma}})]. \quad (131)$$

In this expression  $a$  is the diameter of the spheres and  $\mathbf{b}_{\hat{\boldsymbol{\sigma}}}(1, 2)$  is an operator that changes postcollisional into precollisional velocities in all arguments of the function it acts upon, or

$$\mathbf{b}_{\hat{\boldsymbol{\sigma}}}(1, 2) f(x_1, x_2 \cdots x_N) = f(x'_1, x'_2 \cdots x_N), \quad (132)$$

with  $\mathbf{r}'_1 = \mathbf{r}_1$ ;  $\mathbf{r}'_2 = \mathbf{r}_2$ ;  $\mathbf{v}'_1 = \mathbf{v}_1 - \hat{\boldsymbol{\sigma}}(\hat{\boldsymbol{\sigma}} \cdot \mathbf{v}_{12})$ ;  $\mathbf{v}'_2 = \mathbf{v}_2 + \hat{\boldsymbol{\sigma}}(\hat{\boldsymbol{\sigma}} \cdot \mathbf{v}_{12})$  and all other coordinates remaining unchanged. The subscript - in (131) is needed because the binary collision operators for forward and backward time evolution of distribution functions are slightly different (the same applies to the binary collision operator defined above for smooth potentials). In addition one has forward and backward unbarred binary collision operators describing the time evolution of phase functions. For more details on this, see [6]. Notice that the hard-sphere binary collision operator may be obtained from the binary collision operator for a smooth potential, in the limit that this approaches the hard-sphere potential (for example: the potential  $\phi(r) = \phi_0(r/a)^{-\gamma}$  in the limit  $\gamma \rightarrow \infty$ ).

The first hierarchy equation for hard spheres follows from (130) as

$$\begin{aligned} & \frac{\partial}{\partial t} f(\mathbf{r}_1, \mathbf{v}_1, t) + \nabla_{\mathbf{r}_1} \cdot (\mathbf{v}_1 f(\mathbf{r}_1, \mathbf{v}_1, t)) + \nabla_{\mathbf{v}_1} \cdot (\mathbf{F} f(\mathbf{r}_1, \mathbf{v}_1, t)) = \\ & = \int d\mathbf{r}_2 \int d\mathbf{v}_2 \bar{\mathbf{T}}_-^{(hs)}(1, 2) F_2(\mathbf{r}_1, \mathbf{v}_1, \mathbf{r}_2, \mathbf{v}_2, t). \end{aligned} \quad (133)$$

It is tempting to think of a closure approximation for this equation. For general density Boltzmann's Stosszahlansatz cannot be applied, because even in equilibrium it does not hold. In fact, in equilibrium the pair distribution assumes the form

$$F_2^{eq}(x_1, x_2) = f^{eq}(x_1) f^{eq}(x_2) g(\mathbf{r}_1, \mathbf{r}_2 | n(\mathbf{r})). \quad (134)$$

Here  $g$  is the equilibrium pair correlation function, which is a functional of the density field (remember that in equilibrium an almost arbitrary density

field may be imposed by adjusting the external potential). This relation suggests a generalization of the Stosszahlansatz to

$$\Theta(-\mathbf{v}_{12} \cdot \mathbf{r}_{12})F_2(x_1, x_2, t) = \Theta(-\mathbf{v}_{12} \cdot \mathbf{r}_{12})f(x_1, t)f(x_2, t)g(\mathbf{r}_1, \mathbf{r}_2|n(\mathbf{r}, t)). \quad (135)$$

This closes the first hierarchy equation, since  $n(\mathbf{r}, t)$  follows from  $f(\mathbf{r}, t)$  by integration over the velocity. The resulting equation is known as the revised Enskog equation. It yields the correct form of the hierarchy equation in equilibrium and therefore will at least allow for the correct equilibrium form of  $f$  as a stationary solution. We will see that in fact this solution is unique, as one can obtain an  $H$ -theorem for the revised Enskog equation. In (135) the  $\Theta$  function has been inserted to stress that the factorization of the velocity dependence, which is still akin to the Stosszahlansatz only is supposed to hold for precollisional configurations. Like at low densities it is inconsistent having this property for both precollisional and postcollisional configurations, except in equilibrium.

At low densities the pair correlation function approaches unity, hence the revised Enskog equation approaches to the Boltzmann equation. But notice that besides through the factor  $g$  it differs from this in that both single particle distribution functions are evaluated at the actual positions of the particles and not at the same position. For low density this difference usually becomes unimportant, because this difference in position becomes small compared to the mean free path. But in special situations, like near boundaries or shock fronts the difference may be relevant.

In fact Enskog[7] was the first one to propose a kinetic equation that may be considered as based on a closure like (135). However, for the pair correlation function of the colliding spheres he did not propose the functional form given here, but instead the pair correlation function of a system at a uniform density, chosen as the density at the midpoint between the colliding particles. Due to reflection symmetry one can show[9] that for a one component system the two expressions are the same, up to corrections of order  $(\nabla n)^2$ . Therefore they give rise to the same form of the Navier-Stokes equations. For hard-sphere mixtures though, the pair correlation functions between different species lack this reflection symmetry, so there is no obvious choice for the point at which the densities should be chosen[8]. It turns out that in this case the revised Enskog equation yields transport coefficients satisfying all required Onsager symmetries, whereas generalizations along the lines of Enskog do not[10, 9].

## 4 Enskog and Revised Enskog Equation

In this section we will investigate the Enskog and revised Enskog equations in some detail. We will mostly consider the latter, since it obviously is to be preferred, but occasionally will remark on where it really differs from the original Enskog equation. First of all, let's write the revised Enskog equation in the presence of a boundary collision term. It becomes:

$$\begin{aligned} \frac{\partial}{\partial t} f(\mathbf{r}_1, \mathbf{v}_1, t) + \nabla_{\mathbf{r}_1} \cdot (\mathbf{v}_1 f(\mathbf{r}_1, \mathbf{v}_1, t)) + \nabla_{\mathbf{v}_1} \cdot (\mathbf{F}(\mathbf{r}_1) f(\mathbf{r}_1, \mathbf{v}_1, t)) = \\ = \int d\mathbf{r}_2 \int d\mathbf{v}_2 \bar{\mathbf{T}}_-(1, 2) g(\mathbf{r}_1, \mathbf{r}_2 | n(\mathbf{r}, t)) f(\mathbf{r}_1, \mathbf{v}_1, t), f(\mathbf{r}_2, \mathbf{v}_2, t) + \\ \bar{\mathbf{T}}_{W-} f(\mathbf{r}_1, \mathbf{v}_1, t), \end{aligned} \quad (136)$$

where  $\bar{\mathbf{T}}_-$  was defined in (131) and the superscript *hs* was left out.

### 4.1 Maximum entropy ensembles

Although the revised Enskog equation is not exact, it may be obtained from the BBGKY hierarchy as a first approximation in a few different ways. One way is to obtain it as the short-time limit of a formally exact kinetic equation for hard sphere systems[10, 11] of the form

$$(\partial/\partial t + \mathbf{v} \cdot \nabla_{\mathbf{r}} + \mathbf{F}(\mathbf{r}) \cdot \nabla_{\mathbf{v}}) f(\mathbf{r}, \mathbf{v}, t) = \int_0^t d\tau \bar{\mathbf{B}}(\tau) f(\mathbf{r}, \mathbf{v}, t - \tau)$$

, where the collision operator  $\bar{\mathbf{B}}$  is approximated by its instantaneous contribution. To obtain better approximations one has to analyze the structure of the collision operator in careful detail. In Ref.[10] this was done by examining a digrammatic expansion in terms of graphical elements representing binary collision operators, Mayer cluster functions and free streaming operators. Alternatively Bławdziewicz and Cichocki[11] developed an algebraic representation of this operator. From both approaches it appears there is no obvious systematic way to go beyond the Enskog approximation, but, guided by physical insight and results from computer simulations people have managed to identify the most important terms beyond the Enskog operator, at least for not too high densities.

An alternative approach relies on finding closures of the BBGKY hierarchy based on the principle of maximum entropy ensembles [12, 13, 14]. The

general idea of this approach is to specify the  $n$ -particle distribution function (and thereby also all lower order distribution functions) plus the average energy in case this is not fixed already by the  $n$ -particle distribution function, and then find the  $N$ -particle ensemble that maximizes entropy under the constraint that it reproduces the prescribed  $n$ -particle distribution function correctly. This ensemble will be a functional of the  $n$ -particle distribution function and the same will therefore hold for the  $n + 1$ -particle distribution function, so the hierarchy is closed indeed. Since we know that isolated systems always evolve towards a state of higher entropy and may assume that near equilibrium the  $n$ -particle distribution functions are determined by the hydrodynamic densities, which change only on macroscopic time scales, we can expect that the maximum entropy ensembles constructed this way are pretty good approximations to the actual nonequilibrium ensembles that would represent the system.

Performing this closure at the level of the single particle distribution one finds that the maximal entropy ensemble is of the form[15, 16]

$$\rho_N^{ME}(\Gamma, t) = \frac{1}{Z_N(t)N!h^{dN}} W_N(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) \exp \left[ - \sum_{i=1}^N \lambda(x_i, t) \right], \quad (137)$$

where  $h$  is Planck's constant and  $Z_N(t)$  is a time dependent partition function, fixed, as usual by the normalization of  $\rho_N^{ME}(\Gamma, t)$ . Further,  $W_N$  is the so-called overlap function, which vanishes for all configurations where at least two spheres overlap each other and is unity for all remaining, non-overlapping configurations. A specific form for hard spheres is

$$W_N(\Gamma) = \prod_{i < j \leq N} \Theta(r_{ij} - a). \quad (138)$$

One easily sees that the spatial correlation functions resulting from this ensemble are independent of the velocity dependence of  $\lambda(x, t)$ , so they are the same as for an equilibrium ensemble with the same density field. Consequently the two-particle distribution function for this ensemble is precisely the one used for the closure giving rise to the revised Enskog equation<sup>16</sup>. We

---

<sup>16</sup>Notice this closure implies complete absence of correlations between particle velocities for all values of  $x_1$  and  $x_2$ , including postcollisional ones. As argued before, this is not valid outside of equilibrium, so the maximum entropy ensemble cannot give a completely correct description at all times.

will see in the next subsection that one can derive an  $H$ -theorem for the revised Enskog equation based on the closure through the maximum entropy ensemble.

Closures using the maximum entropy principle at higher values of  $n$  have also been reported in the literature[18].  $H$ -theorems have been derived for all of these. However, the kinetic equations resulting from these closures are forbiddingly hard to solve, already for  $n = 2$ , even though they do contain important memory effects that are missing in the Enskog equations.

## 4.2 $H$ -theorem for the Revised Enskog Equation

The first issue to be dealt with for the derivation of an  $H$ -theorem is how to define an  $H$ -function at arbitrary density. Here we can let ourselves be guided by the identification of  $-k_B H$  with the entropy. According to the maximum entropy procedure[12, 13, 15, 14] the  $H$ -function then must be defined as

$$H(t) = \int d\Gamma \rho_N^{ME}(\Gamma, t) \ln \rho_N^{ME}(\Gamma, t) \quad (139)$$

Now, if the time evolution of  $\rho_N^{ME}(\Gamma, t)$  were given by Liouville's equation the  $H$ -function would remain strictly constant as a consequence of Liouville's theorem, as noted already by Gibbs. This by itself almost suffices to prove an  $H$ -theorem. Suppose a system at initial time  $t_0$  is described correctly by the ensemble  $\rho_N^{ME}$ . Then the time evolution of  $f(\mathbf{r}, \mathbf{v}, t)$  for this system for an infinitesimal time  $\delta t$  is described correctly by the revised Enskog equation. The actual ensemble describing the system at time  $t_0 + \delta t$  has the same Gibbs entropy as at time  $t_0$ , where it coincides with the maximum entropy expression. But the maximum entropy expression at time  $t_0 + \delta t$  cannot be smaller than the Gibbs entropy, since both ensembles share the same  $f(\mathbf{r}, \mathbf{v}, t_0 + \delta t)$ , hence  $H(t_0 + \delta t) \leq H(t_0)$ . However, this argument does not show that the maximum entropy expression is strictly increasing, unless the system is in equilibrium. For this we need a more technical proof, which is due to Résibois[17]. We reproduce it below in slightly more condensed form.

As a first preparation we introduce the field

$$\alpha(\mathbf{r}_1, t) = \ln \int dx_2 dx_3 \cdots dx_N \frac{1}{Z_N(t) h^{d(N-1)} (N-1)!} \exp \left[ - \sum_{i=2}^N \lambda(x_i, t) W_N(\Gamma) \right], \quad (140)$$

so that

$$\lambda(x_1, t) = \alpha(\mathbf{r}_1, t) - \ln [h^d f_1(x_1, t)]. \quad (141)$$

Equations (137), (139), and (140) allow us to express the time derivative of the  $H$ -function as

$$\begin{aligned} \frac{\partial H}{\partial t} &= - \int d\Gamma \left[ \sum_{i=1}^N \lambda(x_i, t) + \ln(Zh^{dN}) \right] \frac{\partial \rho_M(\Gamma, t)}{\partial t}, \\ &= -N \int dx \lambda(x, t) \frac{\partial}{\partial t} \int d\Gamma \rho_M(\Gamma, t) \delta(x - x_1) \\ &= - \int dx (\alpha(\mathbf{r}, t) - \ln [h^d f_1(x_1, t)]) \frac{\partial f_1(x, t)}{\partial t}, \end{aligned} \quad (142)$$

where we used that the integral over  $\rho$  remains normalized to unity. We consider the two terms on the right-hand side of this equation separately. From Eq. (136) and the definition given in Eq. (140) of  $\alpha$  we obtain

$$\int dx \alpha(\mathbf{r}, t) \frac{\partial f_1(x, t)}{\partial t} = \frac{1}{2} \int dx_1 dx_2 F_2(x_1, x_2, t) \hat{\mathbf{r}}_{12} \cdot \mathbf{v}_{12} \delta(r_{12} - a^+), \quad (143)$$

where  $a^+$  indicates the pair correlation has to be evaluated in the limit of  $r_{12}$  approaching the hard sphere diameter from the outside. Here we used that the term resulting from  $\mathbf{F}(\mathbf{r}_1) \cdot \nabla_{\mathbf{v}_1}$  in Eq. (136) vanishes thanks to the property  $\nabla_{\mathbf{v}_1} \cdot \mathbf{F}(\mathbf{r}_1) = 0$ . Similarly, the contribution resulting from the right-hand side of Eq. (136) vanishes, which can be seen most easily by rewriting it in terms of the adjoint of the binary collision operator (see below). Also the contribution from the wall term vanishes, as a consequence of Eqs. (29-31) and the independence of  $\alpha(x_1, t)$  from  $\mathbf{v}_1$ . For the contribution, resulting from the  $\mathbf{v}_1 \cdot \nabla_{\mathbf{r}_1}$  term in Eq. (136) one has to apply a partial integration with respect to  $\mathbf{r}_1$ . Doing this one should notice that the only dependence of  $\alpha$  on  $\mathbf{r}_1$  occurs through the function  $W_N$ , defined in (138). Rewriting  $\nabla_{\mathbf{r}_1} \alpha$  as

$$\nabla_{\mathbf{r}_1} \alpha = \frac{h^{-d} \exp -\lambda(x_1, t) \nabla_{\mathbf{r}_1} e^\alpha}{h^{-d} \exp -\lambda(x_1, t) e^\alpha},$$

and using Eqs. (141) and (138) to evaluate the numerator one easily arrives at the resulting Eq. (143).

For the second term in Eq. (142) both free streaming contributions vanish on integration. The term resulting from the right-hand side of Eq. (136) may

be rewritten as

$$\begin{aligned}
& \frac{1}{2} \int dx_1 \ln \left[ \frac{f_1(x_1, t)}{h^d} \right] \bar{\mathbf{T}}_-(1, 2) F_2(x_1, x_2, t) \\
&= \frac{1}{2} \int dx_1 dx_2 F_2(x_1, x_2, t) \mathbf{T}_+(1, 2) \ln \left[ \frac{f_1(x_1, t) f_1(x_2, t)}{h^{2d}} \right] \\
&= -\frac{1}{2} \int dx_1 dx_2 F_2(x_1, x_2, t) \mathbf{v}_{12} \cdot \hat{\mathbf{r}}_{12} \delta(r_{12} - a^+) \\
&\quad \theta(-\mathbf{v}_{12} \cdot \hat{\mathbf{r}}_{12}) \ln \frac{f_1(x_1, t) f_1(x_2, t)}{f_1(x'_1, t) f_1(x'_2, t)}, \tag{144}
\end{aligned}$$

where the primes indicate again the replacement of the velocities by their restituting values and  $\mathbf{T}_+(1, 2)$  is the adjoint of  $\bar{\mathbf{T}}_-(1, 2)$ , given by

$$\mathbf{T}_+(1, 2) = a^{d-1} \int_{\mathbf{v}_{12} \cdot \hat{\boldsymbol{\sigma}} < 0} d\hat{\boldsymbol{\sigma}} |\mathbf{v}_{12} \cdot \hat{\boldsymbol{\sigma}}| \delta(a\hat{\boldsymbol{\sigma}} - \mathbf{r}_{12}) (\mathbf{b}_{\hat{\boldsymbol{\sigma}}}(1, 2) - 1), \tag{145}$$

The wall collision term finally yields the contribution

$$\left( \frac{\partial H}{\partial t} \right)_W = \int dx_1 \ln \left[ \frac{f_1(x_1, t)}{h^d} \right] \bar{\mathbf{T}}_W f_1(x_1, t). \tag{146}$$

Collecting results one finds

$$\begin{aligned}
\frac{\partial H}{\partial t} &= -\frac{1}{2} \int dx_1 dx_2 F_2(x_1, x_2, t) \mathbf{v}_{12} \cdot \hat{\mathbf{r}}_{12} \delta(r_{12} - a^+) \\
&\quad \left[ 1 - \theta(-\mathbf{v}_{12} \cdot \hat{\mathbf{r}}_{12}) \ln \frac{f_1(x_1, t) f_1(x_2, t)}{f_1(x'_1, t) f_1(x'_2, t)} \right] \\
&\quad + \int dx_1 \ln \left[ \frac{f_1(x_1, t)}{h^d} \right] \bar{\mathbf{T}}_W f_1(x_1, t). \tag{147}
\end{aligned}$$

Next we use in this equation that  $\ln y/z \leq 1 - z/y$ , with  $y = f_1(x_1, t) f_1(x_2, t)$  and  $z = f_1(x'_1, t) f_1(x'_2, t)$ , with equality holding only if  $y = z$ . Thus,

$$\begin{aligned}
\frac{\partial H}{\partial t} &\leq -\frac{1}{2} \int dx_1 dx_2 \mathbf{v}_{12} \cdot \hat{\mathbf{r}}_{12} \delta(r_{12} - a^+) [1 - \theta(-\mathbf{v}_{12} \cdot \hat{\mathbf{r}}_{12}) (1 - b_{\hat{\mathbf{r}}_{12}}(1, 2))] \\
&\quad F_2(x_1, x_2, t) + \int dx_1 \ln \left[ \frac{f_1(x_1, t)}{h^d} \right] \bar{\mathbf{T}}_W f_1(x_1, t). \tag{148}
\end{aligned}$$

Letting the  $b_{\hat{r}_{12}}$ -operator act to the left one easily finds that the first term on the right-hand side vanishes. Hence the final form of the H-theorem for the Revised Enskog equation becomes

$$\frac{\partial H}{\partial t} \leq \int dx_1 \ln \left[ \frac{f_1(x_1, t)}{h^d} \right] \bar{\mathbf{T}}_W f_1(x_1, t), \quad (149)$$

In case the wall collision operator satisfies the thermostat condition, Eq. (42) the steps of subsection 2.2 may be repeated to arrive at Eq. (63). Therefore all results obtained from that will remain valid, such as the free energy being monotonically non-increasing with time, in case the wall has a uniform temperature.

Let us next consider the conditions under which the inequalities hold as equalities. The inequality leading from (147) to (148) reduces to an equality under the condition

$$\ln f(x_1, t) + \ln f(x_2, t) = \ln f(x'_1, t) + \ln f(x'_2, t). \quad (150)$$

This is satisfied for [17]

$$\ln f(x_i, t) = \frac{1}{2} \beta m |\mathbf{v}_i - \mathbf{u}|^2 + \ln n. \quad (151)$$

We find that, due to the difference in position of the two colliding particles temperature, density and average local velocity have to be equal at the two positions and consequentially over the full volume of the system. Unlike for the Boltzmann equation the time derivative of the  $H$ -function does not vanish for local equilibrium distributions.

We finish this subsection by remarking that Mareschal et al. have derived a local H-theorem for the Revised Enskog equation [19], which has been shortened and simplified later by Piasecki [20]. This is similar to the local H-theorem for the Boltzmann equation, as formulated in Eq. (59). Apparently the form of the local  $h$ -function cannot be defined entirely uniquely, as one can redistribute the local entropy density in any point in somewhat arbitrary ways over the surroundings of this point. In fact a similar problem exists with the definition of a microscopic potential energy (e.g. it may be defined as a sum of  $\delta$ -functions localized at the centers of the interacting particles, but it may also be spread over the line pieces connecting these centers). On macroscopic time and length scales this is inconsequential.

## 5 Mode coupling and long time behavior of hydrodynamics

Kinetic theory can be used to study the long time behavior of correlation functions between hydrodynamic densities, or equivalently the long time characteristics of the hydrodynamic decay of systems that are close to equilibrium[22]. In this limit the kinetic equations can be shown to reduce to *mode coupling equations*. These describe the slow decay of hydrodynamic currents, due to couplings to products of hydrodynamic densities, so the nonlinearities in the hydrodynamic equations are crucial for these effects. An alternative way to obtain the mode coupling equations starts from fluctuating hydrodynamics. That is, to the usual hydrodynamic equations fluctuating Langevin terms are added describing the deviations in individual systems from the average behavior of an ensemble of non-equilibrium systems. To derive the fluctuating hydrodynamic equations from first principles one has to go through an intermediate step of deriving fluctuating kinetic equations, such as a fluctuating Boltzmann equation from the microscopic equations of motion. However, one may postulate the hydrodynamic equations on the basis of conservation laws and empirical observations and add fluctuating terms in order to satisfy fluctuation-dissipation requirements. This provides a shortcut to the mode coupling equations[23]. To simplify the analysis and to allow for a simple study of the behavior under low dimensionality I will first consider the fluctuating Burgers equation.

### 5.1 Fluctuating Burgers equation[21]

Consider driven diffusive systems, described by

$$\frac{\partial c(\mathbf{r}, t)}{\partial t} + \nabla \cdot \mathbf{j}(\mathbf{r}, t) = 0 \quad \text{[continuity equation]}$$

with

$$\mathbf{j}(\mathbf{r}, t) = -D(c)\nabla c(\mathbf{r}, t) + c(\mathbf{r}, t)\mathbf{u}(c(\mathbf{r}, t)) + \tilde{\mathbf{j}}_L(\mathbf{r}, t).$$

Here  $\tilde{\mathbf{j}}_L(\mathbf{r}, t)$  is a Langevin noise term. It is supposed to be gaussian white noise, with average  $\langle \tilde{\mathbf{j}}_L(\mathbf{r}, t) \rangle = 0$  and variance  $\langle \tilde{\mathbf{j}}_L(\mathbf{r}, t) \tilde{\mathbf{j}}_L(\mathbf{r}', t') \rangle = N\delta(\mathbf{r} - \mathbf{r}')\delta(t - t')\mathbf{1}$ . The brackets indicate an average over all realizations of the stochastic time evolution and the noise strength  $N$  will be specified a little

below.. Expand the convective current, with  $c(\mathbf{r}, t) = \bar{c} + \tilde{\phi}(\mathbf{r}, t)$

$$c(\mathbf{r}, t)\mathbf{u}(c(\mathbf{r}, t)) = \bar{c}\mathbf{u}(\bar{c}) + \mathbf{v}\tilde{\phi}(\mathbf{r}, t) + \mathbf{w}\tilde{\phi}^2(\mathbf{r}, t)$$

with

$$\mathbf{v} = \frac{\partial}{\partial c}(c\mathbf{u}(c))_{c=\bar{c}} = \mathbf{u}(\bar{c}) + \bar{c}\left(\frac{\partial\mathbf{u}}{\partial c}\right)_{c=\bar{c}} \quad (152)$$

$$\mathbf{w} = \frac{1}{2}\frac{\partial^2}{\partial c^2}(c\mathbf{u}(c)) = \frac{1}{2}\frac{\partial\mathbf{v}}{\partial c} \quad (153)$$

both evaluated at  $c = \bar{c}$ .

$D$  may be approximated by  $D(\bar{c})$ .

$$\longrightarrow \left(\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla\right)\tilde{\phi}(\mathbf{r}, t) = +D\nabla^2\tilde{\phi}(\mathbf{r}, t) - \mathbf{w} \cdot \nabla\tilde{\phi}^2(\mathbf{r}, t) - \nabla \cdot \tilde{\mathbf{j}}_L(\mathbf{r}, t)$$

Next make Galilei transformation:

$$\begin{aligned} \phi(\mathbf{r}, t) &= \tilde{\phi}(\mathbf{r} + \mathbf{v}t, t) && \text{[notice sign!]} \\ \Rightarrow \frac{\partial\phi(\mathbf{r}, t)}{\partial t} &= -\mathbf{w} \cdot \nabla\tilde{\phi}^2(\mathbf{r}, t) + D\nabla^2\tilde{\phi}(\mathbf{r}, t) - \nabla \cdot \mathbf{j}_L(\mathbf{r}, t) && (154) \\ &\text{with } \mathbf{j}_L^\ell(\mathbf{r}, t) = \tilde{\mathbf{j}}_L(\mathbf{r} + \mathbf{v}t, t) \end{aligned}$$

For Fourier components:

$$\frac{\partial}{\partial t}\hat{\phi}(\mathbf{k}, t) = -\frac{i\mathbf{k} \cdot \mathbf{w}}{V} \sum_{\mathbf{q}} \hat{\phi}(\mathbf{q}, t)\hat{\phi}(\mathbf{k} - \mathbf{q}, t) - Dk^2\hat{\phi}(\mathbf{k}, t) - i\mathbf{k} \cdot \mathbf{j}_L(\mathbf{k}, t) \quad (155)$$

Can be transformed to

$$\begin{aligned} \hat{\phi}(\mathbf{k}, t) &= e^{-Dk^2(t-t_0)}\hat{\phi}(\mathbf{k}, t_0) - \int_0^{t-t_0} d\tau \left[ \frac{i\mathbf{k} \cdot \mathbf{w}}{V} e^{-Dk^2\tau} \sum_{\mathbf{q}} \hat{\phi}(\mathbf{q}, t-\tau)\hat{\phi}(\mathbf{k} - \mathbf{q}, t-\tau) \right. \\ &\quad \left. - e^{-Dk^2\tau} i\mathbf{k} \cdot \mathbf{j}_L(\mathbf{k}, t-\tau) \right] \quad (156) \end{aligned}$$

Iterate this equation and make a diagrammatic expansion for the time correlation function  $\langle\phi(\hat{\boldsymbol{\sigma}}, t), \phi(\hat{\boldsymbol{\sigma}}, t_0)\rangle$  with line pieces representing factors  $E^{-Dq_j^2\tau_j}$  and vertices factors  $\frac{i\mathbf{q}_j \cdot \mathbf{w}}{V} \int_0^{t-t_j} dt_{j+1} \sum_{\mathbf{q}_{j+1}}$ . In addition factors  $\hat{\phi}(\mathbf{k}, t)$  or

$\hat{\phi}(k, t_0)$  will be denoted by points and factors  $i\mathbf{k} \cdot \mathbf{j}_L(\mathbf{k}, t - \tau)$  by asterisks. The first iteration of

$$\hat{S}(\mathbf{k}, t) = \frac{1}{V} \langle \hat{\phi}(-\mathbf{k}, t) \hat{\phi}(\mathbf{k}, 0) \rangle \quad (157)$$

$$= \frac{1}{V} \langle \hat{\phi}(-\mathbf{k}, t_0 + t) \hat{\phi}(\mathbf{k}, t_0) \rangle \quad (158)$$

Note that, to leading order,

$$e^{-Dk^2 t} = \frac{\hat{S}(\mathbf{k}, t)}{\hat{S}(\mathbf{k})}$$

with  $\hat{S}(\mathbf{k}) = \hat{S}(\mathbf{k}, 0)$ . After resummation this becomes

$$= \frac{\hat{S}(\mathbf{k}, t)}{\hat{S}(\mathbf{k})}$$

Leading order mode-coupling from one-loop diagrams

$$-i\mathbf{k} \cdot \mathbf{w} \quad -i\mathbf{q} \cdot \mathbf{w} + i(\mathbf{k} - \mathbf{q}) \cdot \mathbf{w} = -i\mathbf{k} \cdot \mathbf{w}$$

$$\Rightarrow \frac{\partial}{\partial t} \hat{S}(\mathbf{k}, t) = -Dk^2 \hat{S}(\mathbf{k}, t) - 2 \frac{(\mathbf{w} \cdot \mathbf{k})^2}{\hat{S}(\mathbf{k}) V} \int_0^{t-t_0} d\tau \sum_{\mathbf{q}} \hat{S}(\mathbf{q}, \tau) \hat{S}(\mathbf{k} - \mathbf{q}, \tau) \hat{S}(\mathbf{k}, t - \tau) \quad (159)$$

from

$$+ \frac{(\mathbf{w} \cdot \mathbf{k})^2}{\hat{S}(\mathbf{k}) V} \int_0^{t-t_0} d\tau \sum_{\mathbf{q}} \hat{S}(\mathbf{q}, \tau) \hat{S}(\mathbf{k} - \mathbf{q}, \tau) \hat{S}(\mathbf{k}, t - \tau)$$

Note that here  $\hat{S}(\mathbf{k})$ ,  $\hat{S}(\mathbf{q})$  and  $\hat{S}(\mathbf{k} - \mathbf{q})$  have all been identified with  $\lim_{k \rightarrow 0} \hat{S}(\mathbf{k})$  and therefore with each other. Justified for long times. Further, in

the limit  $V \rightarrow \infty$ ,  $\frac{1}{V} \sum_{\mathbf{q}}$  approaches  $\frac{1}{(2\pi)^d} \int d\mathbf{q}$ .

The Green-Kubo formalism relates the time derivative of  $\hat{S}(\mathbf{k}, t)$  to a current-current time correlation function (see e.g.[24]) through

$$\frac{\partial}{\partial t} \hat{S}(\mathbf{k}, t) = -k^2 \int_0^\infty d\tau \hat{M}(\mathbf{k}, \tau) \hat{S}(\mathbf{k}, t - \tau) \quad (160)$$

with

$$\begin{aligned} \lim_{k \rightarrow 0} \hat{M}(\mathbf{k}, \tau) &= \hat{\mathbf{k}} \cdot \langle (\mathbf{J}(0) - \langle \mathbf{J}(0) \rangle - \frac{\partial J(\bar{c})}{\partial \bar{c}} (N(0) - \langle N \rangle) \\ &(\mathbf{J}(t) - \langle \mathbf{J} \rangle - \frac{\partial J(\bar{c})}{\partial \bar{c}} (N(t) - \langle N \rangle)) \rangle \end{aligned} \quad (161)$$

Approximating  $\hat{S}(\mathbf{k}, t)$  by  $\hat{S}(0)e^{-Dk^2 t}$  one obtains for  $d > 2$  from (159)

$$\hat{M}(\hat{\mathbf{k}}, t) = \frac{2(\mathbf{w} \cdot \hat{\mathbf{k}})^2 \hat{S}(0)}{(8\pi Dt)^{d/2}} \quad [\text{differs from } C \text{ in [21]}]$$

For  $d = 1, 2$  the mode coupling terms dominate the diffusion equation.

To analyze this for  $d = 1$ , first introduce dimensionless variables:  $\tau = \alpha t$ ;  $\kappa = \beta k$ ;  $\Sigma(\kappa, \tau) = \frac{\hat{S}(\frac{\kappa}{\beta}, \frac{\tau}{\alpha})}{\hat{S}(0)}$  with  $\alpha = \frac{w^4 S^2(0)}{8D^3}$   $\beta = \frac{4D^2}{w^2 \hat{S}(0)} \Rightarrow$

$$\frac{\partial \Sigma(\kappa, \tau)}{\partial \tau} = -\frac{1}{2} \kappa^2 \left[ \Sigma(\kappa, \tau) + \frac{2}{\pi} \int_{-\infty}^{\tau} d\sigma \int_{-\infty}^{\infty} d\lambda \Sigma(\lambda, \sigma) \Sigma(\kappa - \lambda, \sigma) \Sigma(\kappa, \tau - \sigma) \right] \quad (162)$$

In the limit  $\kappa \rightarrow 0$   $\tau \rightarrow \infty$  one may look for a solution of the form  $\Sigma(\kappa, \tau) = h(\kappa \tau^{2/3})$  because the mode-coupling term scales as  $\kappa^3 \tau^2 \Sigma^2$  in comparison to  $\frac{\partial \Sigma}{\partial \tau}$ . The diffusive term scales as  $\kappa^2 \tau$ , which is  $\sim \tau^{-1/3}$  under scaling of  $h$ . So it becomes small for large  $\tau$  indeed. Inserting the scaling form into (162) one may rewrite this as

$$\frac{dh(x)}{dx} = -\frac{9}{4\pi} x \int_{-\infty}^1 ds s^{-1/2} \int_{-\infty}^{\infty} dy h(y) h(sx - y) h(x(1 - s^{3/2})^{2/3})$$

From this scaling it follows that

$$\frac{\partial^2}{\partial k^2} \log \hat{S}(k, t)_{k \rightarrow 0} = \frac{\partial^2}{\partial k^2} \log [h(kt^{2/3})] = \frac{t^{4/3} h''(0)}{h(0)},$$

where  $h'(0) = 0$  has been used. On the other hand,

$$\begin{aligned}
h(\mathbf{k}, t) &= \frac{1}{V} \left\langle \sum_{j\ell} e^{i\mathbf{k}\cdot\mathbf{r}_j(t)} e^{-i\mathbf{k}\cdot\mathbf{r}_\ell(0)} \right\rangle & (163) \\
&= -\frac{1}{V} \left\langle \sum_{j\ell} (1 + i\mathbf{k} \cdot (\mathbf{r}_j(t) - \mathbf{r}_\ell(0)) - \frac{1}{2}\mathbf{k}\mathbf{k} : (\mathbf{r}_j(t) - \mathbf{r}_\ell(0))(\mathbf{r}_j(t) - \mathbf{r}_\ell(0))) \right\rangle \\
&\Rightarrow \frac{\partial^2}{\partial k^2} \log \hat{S}(k, t) = \frac{N^2 \langle [(\mathbf{R}(t) - \mathbf{R}(0)) \cdot \hat{\mathbf{k}}]^2 \rangle}{N^2}
\end{aligned}$$

So the mean square displacement of the center of mass increases as  $t^{4/3}$  in the comoving frame. Its second time derivative is the current-current time correlation function in the comoving frame, behaving as  $t^{-2/3}$ .

This may be tested on the Asymmetric Simple Exclusion Process [ASEP]. This model consists of a one-dimensional lattice, each site of which may either be empty or occupied by a single particle. The particles may jump to unoccupied neighboring sites with jump rates defined as follows:



$$\begin{aligned}
\text{Jump rate to unoccupied site} &= p\Gamma && \text{to right} \\
&= (1-p)\Gamma && \text{to left}
\end{aligned}$$

This is called asymmetric for  $p \neq 1/2$ . The stationary distribution of particle configurations gives equal weight to all allowed configurations (then both the gain and the loss rate for a configuration equals  $\Gamma n_{clusters}$ ). The average current in the stationary state equals

$$c(1-c)(2p-1)\Gamma$$



The contributions to this from jumps to the right and to the left are  $c(1-c)\Gamma p$  and  $(1-c)c\Gamma(1-p)$  respectively. From this one obtains

$$v = (2p-1)\Gamma(1-2c) \quad (164)$$

$$w = -(2p-1)\Gamma \quad (165)$$

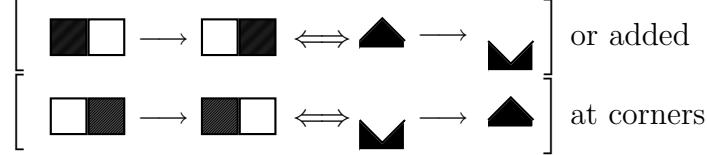
Simulations on this model confirm the  $t^{4/3}$ -behavior of the mean-square displacement of the center of mass.

## 5.2 The Kardar-Parisi-Zhang equations[27]

As an introduction let us consider again the ASEP. It may be considered as a  $1d$  interface model by identifying a particle with a surface element  and an empty space as a surface element .



The dynamics correspond to a growth process where units of mass are removed



The equation describing this is obtained by integrating over  $x : \int_0^x dx' [2\rho(x') - 1] \approx h(x)$ .

The  $1d$  KPZ-equation follows from the  $1d$  fluctuating Burgers equation as

$$\frac{\partial h(x, t)}{\partial t} = D \frac{\partial^2 h(x, t)}{\partial x^2} - w \left( \frac{\partial h(x, t)}{\partial x} \right)^2 + \eta_L(x, t) \quad (166)$$

A first remark is that in a situation of steady growth (or evaporation/solution) one should add a constant term  $v_0$  equal to the average growth speed, on the right hand side. Mathematically this makes no difference. One may describe the process in a comoving frame through  $\tilde{h}(x, t) = h(x, t) - v_0 t$ . This satisfies (166) again.

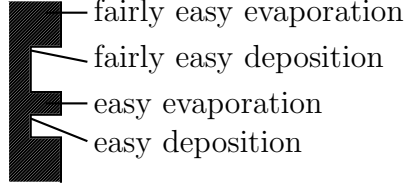
A second remark is that (166) generalizes in  $d$  dimensions (especially  $d = 2$  is physically relevant) to

$$\frac{\partial h(\mathbf{r}, t)}{\partial t} = +D \nabla^2 h(\mathbf{r}, t) - w |\nabla h(\mathbf{r}, t)|^2 + \eta_L(\mathbf{r}, t). \quad (167)$$

Notice that for  $d > 1$  this is not equivalent to the Burgers equation.

A third remark is about the physical interpretation of the various terms. The noise term  $\eta_L$  describes random fluctuations in the deposition and evaporation process. The term  $-D \nabla^2 h$  describes the effects of diffusion of adatoms and vacancies along the surface. Finally the non-linear term  $W |\nabla h|^2$  describes the effect of the surface slope on the average evaporation or growth rate. There are two causes for this:

1. On rough surfaces both evaporation and deposition proceeds more easily, because fewer bonds have to be broken respectively more can be saturated.



2. Particles (or holes) left on a flat surface diffuse easily until they reach some step edge. (so the dynamics tends to enhance their smoothness). Notice that for the ASEP horizontal surfaces are rough and sloped ones are smooth. This implies that the constant  $w$  is negative. For the model properties this makes no difference: on transforming from  $h$  to  $-h$  the relative signs of  $w$  and the other constants are changed.

### 5.3 Scaling behavior

In models of growth dynamics (or different types of dynamics) functions of  $\mathbf{r}$  and  $t$  for large  $r$  and  $t$  often exhibit *scaling dynamics*:  $G(r, t) = b^{-\zeta} G(br, b^z t)$   
 E.g. in  $d > 2$  consider  $\langle h(\mathbf{r}_0, t)h(\mathbf{r}_0 + \mathbf{r}_1, t) \rangle$ , with  $h(\mathbf{r}, 0) = 0$ , so

$$h(\mathbf{k}, t) = \int_0^t d\tau e^{-Dk^2\tau} \eta(t - \tau, \mathbf{k})$$

$$\Rightarrow \langle h(\mathbf{r}_0, t)h(\mathbf{r}_0 + \mathbf{r}_1, t) \rangle = \frac{2}{(2\pi)^d} \int \frac{d\mathbf{k}}{2Dk^2} \langle \eta^2 \rangle e^{-2Dk^2t} e^{i\mathbf{k}\cdot\mathbf{r}_1} \quad (168)$$

$$\sim t^{-(\frac{d}{2}-1)} e^{-\frac{r^2}{8Dt}}. \quad (169)$$

leading to  $z = 2$  and, with  $b^\zeta = b^{-(d-2)}$ , to  $\zeta = 2 - d$ . [For this function! E.g. for  $\langle |h(\mathbf{r}_0 + \mathbf{r}, t) - h(\mathbf{r}_0, t)| \rangle$  you find half this value.]

For KPZ in  $d = 1$  we had  $\hat{S}(k, t) = h(kt^{2/3})\hat{S}(0) \Rightarrow S(x, t) = \tilde{h}(x/t^{2/3}) \Rightarrow z = 3/2$ .

Whereas  $\langle |h(x, t) - h(x', t)| \rangle \sim |x - x'|^{1/2} \Rightarrow \zeta = \frac{1}{2}$ .

From  $G(r^2/t) \sim r^{-\zeta} G((br)^2/b^2t)$  it follows that

$$\begin{aligned} G(k, t) &\sim G(kt^{2/3}) \\ G(\mathbf{r}, t) &\sim G(t^{2/3}/r) \Rightarrow z = 3/2 \end{aligned}$$

whereas for normal diffusion,

$$G(\mathbf{r}, t) \sim G(r^2/t) \Rightarrow z = 2.$$

## 5.4 The polynuclear growth model. Exact results

Prähofer and Spohn[25] managed to solve exactly a specific one-dimensional growth model within the KPZ universality class; the polynuclear growth model. This model consists of line pieces stacked on top of each other. Growth nuclei are created at a constant rate at completely random positions. After creation a nucleus grows with constant speed  $v$  in both directions, thus creating a new line piece on top of an existing one. When two of these line pieces meet their growth stops in this direction. For this model density-density and current-current time correlations can be solved exactly and expressed in terms of scaling functions, which have been tabulated with great precision by the authors. Since the long time behavior within a universality class is the same for all members, provided parameters are identified correctly, the long time behavior of time correlation functions for the fluctuating Burgers equation can be expressed in terms of the Prähofer-Spohn scaling functions[28]. The most striking results are first of all a prediction for the decay of the current-current time correlation function as

$$\frac{1}{L} \langle \tilde{J}(t) \tilde{J}(0) \rangle = \frac{2.1056}{\Gamma_E(1/3)} \left( \frac{8\hat{S}(0)w^2}{t} \right)^{2/3}, \quad (170)$$

with  $\tilde{J}(t) = \hat{J}(0, t) - \langle \hat{J}(0, t) \rangle - \frac{\partial \hat{J}(\bar{c})}{\partial \bar{c}} (N(0, t) - \langle N \rangle)$  and  $\Gamma_E$  denoting Euler's gamma function. Secondly, for the wave number diffusion coefficient, characterizing the decay rate of a sine-wave of wavelength  $k$  in a periodic system, as  $D(k)k^2$  one finds

$$D(k) = \frac{1}{19.444} \sqrt{\frac{8\hat{S}(0)w^2}{|k|}}. \quad (171)$$

One can obtain more detailed results from the Prähofer-Spohn scaling functions, which are discussed in Ref.[25].

## 5.5 Hydrodynamics in one dimension

Hamiltonian systems in one dimension have three global conservation laws, for mass (or number), momentum and energy. Therefore one has three hydrodynamic equations, which on adding fluctuating terms are of the form

$$\begin{aligned}\frac{\partial}{\partial t}n(x, t) &= -\frac{\partial}{\partial x}[n(x, t)u(x, t)]\frac{\partial}{\partial t}g(x, t) \\ \frac{\partial}{\partial x}[g(x, t)] &= -\frac{\partial p(x, t)}{\partial x} + \frac{\partial}{\partial x}\left\{\zeta(n(x, t), T(x, t))\frac{\partial g(x, t)}{\partial x}\right\} + \frac{\partial \sigma^r(x, t)}{\partial x}.\end{aligned}\tag{172}$$

Plus an equation for the energy density. In these expressions  $g(x, t)$  is the momentum density. The pressure  $p(x, t)$  may be expressed as

$$p(x, t) = \left(\frac{\partial p}{\partial n}\right)_e n(x, t) + \left(\frac{\partial p}{\partial e}\right)_n e(x, t),$$

with  $e(x, t)$  the local energy density. Finally,  $\sigma^r(x, t)$  is the *random stress tensor*, usually represented by gaussian white noise, with zero mean and variances dictated by the fluctuation dissipation theorem. After linearization and taking a spatial Fourier transform these equations can be rearranged into

$$\frac{\partial m\hat{n}(k, t)}{\partial t} = -ik\hat{g}(k, t),\tag{173}$$

$$\frac{\partial \hat{g}(k, t)}{\partial t} = -ik\hat{p}(k, t) - \frac{ik}{n_0 m} [\zeta\hat{g}(k, t) + \hat{\sigma}^r(k, t)],\tag{174}$$

$$T_0\frac{\partial \hat{s}(k, t)}{\partial t} = \frac{\lambda}{n_0 T_0} k^2 \hat{T}(k, t).\tag{175}$$

Here  $\hat{s}(k, t)$  is the local entropy density, defined in similar way as the local pressure as a linear combination of particle and energy density. These equations can be diagonalized. One then finds three eigenmodes, called hydrodynamic modes. These are two sound modes<sup>17</sup>  $a_1(k, t)$  and  $a_{-1}(k, t)$  and a heat mode  $a_H(k, t)$ , given respectively, to leading order in  $k$  by

$$a_\sigma(k, t) = \left(\frac{\beta}{2\rho_0}\right)^{1/2} (c_0^{-1}p(k, t) + \sigma g(k, t)),\tag{176}$$

$$a_H(k, t) = \left(\frac{\beta}{n_0 T_0 C_p}\right)^{1/2} (e(k, t) - h_0 n(k, t)).\tag{177}$$

---

<sup>17</sup>I use  $\sigma = \pm 1$  for right respectively left moving sound modes, rather than positive respectively negative frequency, as is conventional.

Here,  $\sigma=\pm 1$ ,  $T_0$  is the equilibrium temperature,  $\beta = (k_B T_0)^{-1}$ ;  $n_0$  the equilibrium number density and  $\rho=nm$ ;  $C_p=T(\partial s/\partial T)_p$  is the specific heat per particle at constant pressure  $p$ , with  $s$  the equilibrium entropy per particle;  $c_0=(\partial p/\partial \rho)_s^{1/2}$  is the adiabatic sound velocity in the limit of zero wave number and  $h_0$  is the equilibrium enthalpy per particle. The allowed values of  $k$  are of the form  $k = \frac{2\pi n}{L}$ . To leading order in  $k$  the hydrodynamic modes are normalized under the inner product  $(f, g) = \frac{1}{L}\langle f^*g \rangle$ , with  $\langle \rangle$  a grand canonical equilibrium average.

The time correlation functions of the hydrodynamic modes satisfy linear equations involving memory kernels, of similar form as Eq. (160), viz.

$$\frac{\partial \hat{S}_\sigma(k, t)}{\partial t} = -i\sigma c_0 k \hat{S}_\sigma(k, t) - k^2 \int_0^t d\tau \hat{M}_\sigma(k, \tau) \hat{S}_\sigma(k, t - \tau), \quad (178)$$

$$\frac{\partial \hat{S}_H(k, t)}{\partial t} = -k^2 \int_0^t d\tau \hat{M}_H(k, \tau) \hat{S}_H(k, t - \tau). \quad (179)$$

Here  $\hat{S}_\sigma(k, t) = (a_\sigma(k, 0), a_\sigma(k, t))$  etc. Like for the fluctuating Burgers equation the memory kernels may be expressed through a diagrammatic mode coupling expansion as a sum of irreducible skeleton diagrams[29]. These consist of propagators, representing stationary density correlation functions  $\hat{S}_\zeta(\ell, t_\alpha)$ , and vertices representing the coupling of one propagator  $\hat{S}(\ell, t_\alpha)$  to two propagators  $\hat{S}_\mu(q, t_{\alpha'})$  and  $\hat{S}_\nu(\ell - q, t_{\alpha''})$ , with coupling strength  $\ell W_\zeta^{\mu\nu}$ . For the long time dynamics only a few of these 27 couplings are important; only couplings to two sound modes of the same sign or to two heat modes may give rise to long-lived perturbations, all other combinations of pairs of modes rapidly die out through oscillations. From EHvL[24] the relevant non-vanishing coupling strengths to leading order in  $k$  can be obtained as<sup>18</sup>

$$W_\sigma^{\sigma'\sigma'} = \frac{\sigma}{2(\rho\beta)^{1/2}c_0} \left( \frac{\partial c_0 n}{\partial n} \right)_s \quad (180)$$

$$W_\sigma^{HH} = \frac{-\sigma(\gamma - 1)n}{2(\rho\beta)^{1/2}C_p} \left( \frac{\partial C_p}{\partial n} \right)_p \quad (181)$$

$$W_H^{\sigma\sigma} = \frac{\sigma k_B^{1/2} c_0}{(2nC_p)^{1/2}}. \quad (182)$$

Notice that  $W_\sigma^{\sigma'\sigma'}$  does not depend on the value of  $\sigma'$ .

---

<sup>18</sup>For obtaining Eq. (181) from the EHvL expression some thermodynamics is required.

Now a central observation is the following: due to the first term on the right-hand side of Eq. (178) the sound-sound correlation functions will have their weights centered around the positions  $x(t) = x(0) \pm c_0 t$ , in other words, these functions will assume the forms  $\hat{S}_\sigma(k, t) = \exp(-i\sigma c_0 k t) \hat{\Sigma}_\sigma(k, t)$ , with  $\hat{\Sigma}_\sigma(k, t)$  to a first approximation real non-oscillating functions. As a consequence the mode coupling contributions to  $\hat{M}^\sigma$  are dominated by those diagrams in which all vertices are of the type  $V_\sigma^{\sigma\sigma}$ . All other contributions for at least some time will oscillate out of phase with the angular frequency  $\sigma c_0 k$  of the sound mode under consideration. The remaining contributions, especially so if described in a coordinate frame comoving at the speed of sound have exactly the same structure as the terms in the mode coupling expansion for the fluctuating Burgers equation[30]; all propagators correspond to the same type of correlation function and all vertices carry the same weight factor  $W$ , in the case of the Burgers equation given by  $W_B = 2w\sqrt{\hat{S}(0)}$ . Therefore, to leading order in time this memory kernel may be expressed in terms of the Pr"ahofer-Spohn scaling functions, like the memory kernel for the fluctuating Burgers equation.

Let us consider the wave number dependent sound damping constant  $\Gamma(k) = 2\tilde{M}_\sigma(k, 0)$  and the sound currents, defined as  $\hat{J}_\sigma(k, t) = \left(\frac{\beta}{2\rho}\right)^{1/2} \sigma \hat{J}_l(k, t) + \frac{1}{c_0} \hat{J}_H(k, t) - \sigma \left(\frac{\partial p}{\partial n}\right)_e$ , where  $\hat{J}_l(k, t)$  and  $\hat{J}_H(k, t)$  are the longitudinal current and the heat current[24], denoted by EHVl as  $J_l$  and  $J_\lambda$  respectively. Eq. (5.28) of Ref.[25] can now be used to obtain the leading small- $k$  behavior of  $\Gamma(k)$  and long time behavior of  $\langle \hat{J}_\sigma(0, 0) \hat{J}_\sigma(0, t) \rangle$  as

$$\Gamma(k) = \frac{16}{19.444} \sqrt{\frac{V_s^2}{|k|}} \quad (183)$$

$$\frac{1}{L} \langle \hat{J}_\sigma(0, t) \hat{J}_\sigma(0, 0) \rangle = \frac{2.1056}{4\sqrt{3}\Gamma_E(1/3)} \left(\frac{V_s^2}{t}\right)^{2/3}. \quad (184)$$

The leading higher order corrections are obtained by replacing in the diagrammatic expansion of the memory kernel just one pair of vertices of type  $V_\sigma^{\sigma\sigma}$  by vertices of type  $V_\sigma^{-\sigma-\sigma}$  or  $V_\sigma^{HH}$ . Note this can only be done by having the new vertices connected by the same pair of propagators. One easily shows that all these terms add contributions proportional to  $|k|^{-1/3}$  to  $\Gamma(k)$  and contributions proportional to  $t^{-7/9}$  to the current-current correlation function. Since there are infinitely many such contributions, there seems

to be no straightforward way of determining the coefficients exactly. However, estimates based on the simplest contributing diagrams can be made[26]. Further corrections obtain from terms with 4, 6,  $\dots$  vertices of type  $V_\sigma^{-\sigma-\sigma}$  or  $V_\sigma^{HH}$ . Each of these appears to be of the form  $Ck^{-\mu}$  for  $\Gamma(k)$  and  $Dt^{-\nu}$  for the current correlation function, with  $C$  and  $D$  constants and  $\mu$  and  $\nu$  of the form  $\mu = 1/3 - \sum_{j=2}^{\infty} m_j(2/3)^j$  and  $\nu = 2/3 + \sum_{j=2}^{\infty} 2n_j(2/3)^j$  respectively, with  $m_j$  and  $n_j$  natural numbers. Again, for each exponent there is an infinity of contributing terms.

The leading long time behavior of  $\hat{S}_H(k, t)$  is determined in similar way by the sum of all contributions to  $\hat{M}^H(k, t)$  where the first and last vertex are of type  $V_H^{\sigma\sigma}$  and all other vertices are of type  $V_\sigma^{\sigma\sigma}$ , all with the same value of  $\sigma$ . These terms do contain an oscillating factor  $\exp(-i\sigma c_0 kt)$ , but these oscillations are much slower than the oscillations in any of the other terms. Since we have to include the contributions to  $\hat{M}^H$  of either sign of  $\sigma$ , we cannot express  $\hat{S}_H$  directly in terms of the Prähofers-Spohn scaling functions, but we can do so immediately for the memory kernel. A simple analysis yields to leading order

$$\hat{M}_H(k, t) = 2 \frac{V_H^2}{V_s} \cos(\sigma c_0 kt) \hat{M}_{PS}(k, V_s t), \quad (185)$$

with  $V_H = |W_H^{\sigma\sigma}|$ . For the  $k$ -dependent heat conduction coefficient and the heat current time correlation function this leads to the expressions

$$\lambda(k) = nC_p D_T(k) = 2nC_p \left(\frac{V_H}{V_s}\right)^2 V_s^{4/3} \frac{2.1056}{4\sqrt{3}(c_0|k|)^{1/3}}, \quad (186)$$

$$\frac{1}{L} \langle \hat{J}_H(0, t) \hat{J}_H(0, 0) \rangle = 2nC_p \left(\frac{V_H}{V_s}\right)^2 \frac{2.1056}{\Gamma_E(1/3)} \left(\frac{V_s^2}{t}\right)^{2/3}. \quad (187)$$

Higher order corrections may be obtained in similar way as for the sound modes.

The analysis presented here clearly shows that for long times the dynamics of 1d hydrodynamic systems to leading order belongs to the KPZ universality class and can be described exactly by means of the Prähofers-Spohn scaling functions. However, the correction terms decay only slightly faster with time and in most cases will not be negligible.

## References

- [1] P. and T. Ehrenfest, *Über zwei bekannte Einwände gegen das Boltzmannsche H-Theorem*, Z. für Physik **8**, 311-314, (1907).
- [2] S. Chapman and T. G. Cowling, *The Mathematical Theory of Non-Uniform Gases* (third ed.), (Cambridge University Press, Cambridge 1990)
- [3] J. H. Ferziger and H. G. Kaper, *Mathematical Theory of Transport Processes in Gases*, (North Holland, Amsterdam 1972)
- [4] J. H. Hirschfelder, C. F. Curtiss and R. B. Bird, *Molecular Theory of Gases and Liquids*, (Wiley, New York 1954)
- [5] J. V. Sengers, in *Recent Advances in Engineering Science*, edited by A. C. Eringen (Gordon and Breach, New York, 1968), p. 153; J. M. Kincaid, S. Pérez and E. G. D. Cohen, Phys. Rev. A **38**, 3628-3633 (1988)
- [6] M. H. Ernst, J. R. Dorfman, W. Hoegy and J. M. J. van Leeuwen. Physica **45**, 127-146 (1969)
- [7] D. Enskog, K. Sven. Vetenskapsakad. Handl. **63**, 4 (1922)
- [8] L. Barajas, L. S. Garcia-Colin, and E. Piña, J. Stat. Phys. **7**, 161 (1973)
- [9] H. van Beijeren and M. H. Ernst, *The modified Enskog equation*, Physica **68** 437-456 (1973); H. van Beijeren and M. H. Ernst, *The modified Enskog equation for mixtures*, Physica **70** 225-242 (1973)
- [10] H. van Beijeren and M. H. Ernst, *Kinetic theory of hard spheres*, J. Stat. Phys. **21**, 125-167 (1979)
- [11] J. Bławdziewicz and B. Cichocki, *Linear kinetic theory of hard-sphere fluids : I. General formalism*, Physica A **127**, 38-71 (1984); B. Cichocki, *Enskog renormalization in linear kinetic theory*, Physica A **142**, 245-272 (1987)
- [12] J. E. Mayer, J. Chem. Phys. **33**, 1484 (1960)
- [13] A. D. McLachlan and R. A. Harris, J. Chem. Phys. **34**, 1451 (1961)

- [14] E. T. Jaynes, in: Statistical Physics, 1962 Brandeis Summer Institute in Theoretical Physics (Benjamin, New York, 1963), Vol. 3
- [15] R. M. Lewis, J. Math. Phys. 8, 1448 (1967)
- [16] J. Karkheck, G. Stell, *Maximization of entropy, kinetic equations, and irreversible thermodynamics*, Phys. Rev. A **25**, 3302-3327 (1982); R'esibois[17] had already postulated the same ensemble for obtaining closure, without referring to the maximum entropy formalism.
- [17] P. Resibois, Phys. Rev. Lett. **40**, 1409 (1978); P. Resibois, *H-theorem for the (modified) Enskog equation*, J. Stat. Phys. **19**, 593-609 (1978)
- [18] J. Bławdziewicz, B. Cichocki, and H. van Beijeren, *H-theorem for a linear kinetic theory*, J. Stat. Phys. **66**, 607-633 (1992); A. Bednorz, B. Cichocki *General H-theorem for hard spheres*, J. Stat. Phys. **114**, 327-360 (2004)
- [19] M. Mareschal, J. Bławdziewicz and J. Piasecki, *Local entropy production from the revised Enskog equation*, Phys. Rev. Lett. **52**, 1169 (1984)
- [20] J. Piasecki, *Local H-theorem for the revised Enskog equation*, J. Stat. Phys. **48**, 1203-1211 (1987)
- [21] H. van Beijeren, R. Kutner, H. Spohn, Phys. Rev. Lett. **54**, 2026 (1985)
- [22] J. R. Dorfman and E. G. D. Cohen, *Velocity Correlation Functions in Two and Three Dimensions*, Phys. Rev. Lett. **25**, 1257 (1970); *Velocity-Correlation Functions in Two and Three Dimensions: Low Density*, Phys. Rev. A **6**, 776 (1972); *Velocity-correlation functions in two and three dimensions. II. Higher density*, *ibid.* **12**, 292 (1975).
- [23] D. Bedeaux, P. Mazur, Physica **73**, 431 (1974); D. Bedeaux, P. Mazur, Physica **75**, 79 (1974)
- [24] M. H. Ernst, E. H. Hauge, J. M. J. van Leeuwen, J. Stat. Phys. **15**, 7 (1976)
- [25] M. Prähofer, H. Spohn, J. Stat. Phys. **115**, 255 (2004)
- [26] H. van Beijeren, to be published

- [27] J. Krug and H. Spohn in *Solids far from equilibrium*, C. Godrèche ed. (Cambridge University Press, Cambridge, 1992).
- [28] H. van Beijeren, *Exact results for anomalous transport in one dimensional Hamiltonian systems* arXiv:1106.3298v2 [cond-mat.stat-mech]
- [29] A. Khuranat, J. Phys. A: Math. Gen. **18**, 2415 (1985); H. van Beijeren, M. H. Ernst, J. Stat. Phys. **21**, 125 (1979) sec. 7
- [30] H. van Beijeren, R. Kutner, H. Spohn, Phys. Rev. Lett. **54**, 2026 (1985); E. Frey, U. C. Täuber, T. Hwa, Phys. Rev. E **53**, 4424 (1996)
- .
- .