March 13, 1996

The Boltzmann-Sinai Ergodic Hypothesis for Hard Ball Systems

Nándor Simányi[1 2]
The Pennsylvania State University, Department of Mathematics
University Park, PA 16802

Domokos Szász[1]
Mathematical Institute of the Hungarian Academy of Sciences
H-1364, Budapest, P. O. B. 127, Hungary

Abstract. We consider the system of N (≥ 2) elastically colliding hard balls with masses $m_1, \dots, m_N$ moving uniformly in the flat unit torus $\mathbb{T}^\nu$, $\nu \geq 3$. It is proved here that the arising billiard flow possesses the K-mixing property for almost every distribution of the masses $m_1, \dots, m_N$.

1. Introduction.

The proper mathematical formulation of Ludwig Boltzmann's ergodic hypothesis, which has incited so much interest and discussion in the last hundred years, is still not clear. For systems of elastic hard balls on a torus, however, Yakov Sinai, in 1963, [Sin(1963)] gave a stronger, and at the same time mathematically rigorous version of Boltzmann's hypothesis: *the system of an arbitrarily fixed number N of identical elastic hard balls moving in theν-torus $\mathbb{T}^\nu$ $(\nu \geq 2)$ is ergodic — of course, on the submanifold of the phase space specified by the trivial conservation laws.* Boltzmann used his ergodic hypothesis when laying down the foundations of statistical physics, and its various forms are still intensively used in modern statistical physics. The meaning of Sinai's hypothesis for the theory of dynamical systems, which was partially based on the physical arguments of Krylov's 1942 thesis (cf. [K(1979)] and its afterword written by Ya. G. Sinai, [Sin(1979)]), is stressed by the fact that the interaction of elastic hard balls defines the only physical system of an arbitrary number of particles in arbitrary dimension whose dynamical behaviour has been so far at least guessed — except for the completely integrable system of harmonic oscillators. (As to the history of Boltzmann's hypothesis, see the recent work [Sz(1996)].)

---

[1] Research supported by the Hungarian National Foundation for Scientific Research, grants No. 7275 and 16425

[2] On leave from the Mathematical Institute of the Hungarian Academy of Sciences, H-1364, Budapest, P. O. B. 127, Hungary

The aim of the present paper is to essentially establish the aforementioned hypothesis which we call the *Boltzmann-Sinai Ergodic Hypothesis.* Here *essentially* has two meanings:

(i) on one hand, in our main theorem, the K-property is only claimed in dimensions $\nu \geq 3$;

(ii) on the other hand — more generally than in Sinai's formulation — we consider balls with different masses rather than those with necessarily identical masses, though our theroem is only established for mass vectors $\vec{m} := (m_1, \dots, m_N) \in \mathbb{R}^N$ not belonging to a countable union of proper analytic submanifolds.

(The equality of the *radii* of the balls is not essential, but, for simplicity, it will be assumed throughout. For certain values of the radii — just think of the case when they are large — the phase space of our system decomposes into a finite union of different connected components, and these connected parts certainly belong to different ergodic components. Now, according to the wisdom of the ergodic hypothesis, these connected components are expected to be just *the* ergodic components of the system, and on each of them the system should also possess the Kolmogorov-mixing property.)

Let us first specify the model and formulate our result.

Assume that, in general, a system of $N(\geq 2)$ balls, identified as $1, 2, \dots, N$, of masses $m_1, \dots, m_N$ and of radius $r > 0$ are given in $\mathbb{T}^\nu$, the ν-dimensional unit torus ($\nu \geq 2$). Denote the phase point of the ith ball by $(q_i, v_i) \in \mathbb{T}^\nu \times \mathbf{R}^\nu$. A priori, the configuration space $\tilde{\mathbf{Q}}$ of the N balls is a subset of $\mathbb{T}^{N\cdot\nu}$: from $\mathbb{T}^{N\cdot\nu}$ we cut out $\binom{N}{2}$ cylindric scatterers:

$$\tilde{C}_{i,j} = \left\{ Q = (q_1, \dots, q_N) \in \mathbb{T}^{N\cdot\nu} : \mid q_i - q_j \mid < 2r \right\}, \tag{1.1}$$

$1 \leq i < j \leq N$, or in other words $\widetilde{\mathbf{Q}} := \mathbb{T}^{N\nu} \setminus \bigcup_{1 \leq i<j \leq N} \tilde{C}_{i,j}$. The energy $H = \frac{1}{2}\sum_1^N m_i v_i^2$ and the total momentum $P = \sum_1^N m_i v_i$ are first integrals of motion. Thus, without loss of generality, we can assume that $H = \frac{1}{2}$, $P = 0$. (If $P \neq 0$, then the system has an additional conditionally periodic or periodic motion.) Now, for these values of H and P, we define our dynamical system.

In earlier works (cf. [S-Sz(1995)-II] and the references therein), where the masses were identical, one could and — to obtain ergodicity — one had to fix a center of mass. For different masses — as observed in [S-W(1989)] — this is generically not possible and we should follow a different approach.

The equivalence relation Ψ over $\widetilde{\mathbf{Q}}$, defined by $Q \sim_\Psi Q^*$ *if and only if there exists an* $a \in \mathbb{T}^\nu$ *such that for every* $i \in [1, N]$ $\quad q_i^* = q_i + a$, allows us to introduce $\mathbf{Q} := \widetilde{\mathbf{Q}}/\Psi$ (the equivalence relation means, in other words, that for the internal coordinates $q_i - q_j = q_i^* - q_j^*$ holds for every $i, j \in [1, N]$). $\mathbf{Q}$, a compact, flat Riemannian manifold with boundary will actually be the configuration space of our system, whereas its phase space will be the ellipsoid bundle $\mathbf{M} := \mathbf{Q} \times \mathcal{E}$, where $\mathcal{E}$ denotes the ellipsoid $\sum_{i=1}^N m_i(dq_i)^2 = 1$, $\quad \sum_{i=1}^N m_i dq_i = 0$. Clearly, $d := \dim \mathbf{Q} = N\nu - \nu$, and $\dim \mathcal{E} = d - 1$. The well-known Liouville measure μ is invariant with respect to the evolution $S^{\mathbb{R}} := \{S^t : \ t \in \mathbb{R}\}$ of our dynamical system defined by elastic collisions of the balls of masses $m_1, \dots, m_N$ and their uniform free motion. Two remarks:

(i) the collision laws for a pair of balls with different masses are well-known from mechanics and will also be reproduced in equation (3.9);

(ii) the dynamics can, indeed, be defined for $\mu-$a. e. phase point, see the corresponding references in Section 2.

The dynamical system $(\mathbf{M}, S^{\mathbb{R}}, \mu)_{\vec{m}}$ is called *the standard billiard ball system with mass vector* $\vec{m} \in \mathbb{R}_+^N$.

Denote by $R_0 := R_0(N, \nu)$ the interval of those values of $r > 0$, for which the interior $\text{Int}\mathbf{M} = \text{Int}\mathbf{Q} \times \mathcal{E}$ of the phase space of the standard billiard ball flow is connected.

The basic result of our paper is the following

Main Theorem. *For $N \geq 2, \nu \geq 3$ and $r \in R_0$, the standard billiard ball system $(\mathbf{M}, \{S^{\mathbb{R}}\}, \mu)_{\vec{m}}$ is a K-flow apart from a countable union of proper analytic submanifolds of mass vectors $\vec{m} \in \mathbb{R}_+^N$.*

Note 1. For $N \geq 2, \nu = 2$ and $r \in R_0$, our methods provide that the standard billiard ball system $(\mathbf{M}, \{S^{\mathbb{R}}\}, \mu)_{\vec{m}}$ has open ergodic components (and thus of positive measure), and, moreover, on each of them the flow has the K-property apart from a countable union of proper analytic submanifolds of mass vectors $\vec{m} \in \mathbb{R}_+^N$ (see the Remark at the end of section 5).

Note 2. In general, if $r \notin R_0$, then $\text{Int}\mathbf{M}$ decomposes into a finite number of connected components. Then our methods provide that the standard billiard ball flow is a K-system on each of these connected components.

Note 3. As it will be seen in Section 2 (see Lemma 2.1) our system is isomorphic to a *semi-dispersing billiard*, and, what is more, is a semi-dispersing billiard in a weakly generalized sense, and, consequently, we can use the theory of semi-dispersing billiards. As it has been proved in recent manuscripts by N. I. Chernov and C. Haskell, [Ch-H(1995)] on one hand, and by D. Ornstein and B. Weiss, [O-W(1996)] on the other hand, the K-mixing property of a semi-dispersing billiard flow actually implies its Bernoulli property, as well.

Note 4. As far as we know, apart from the theorem of [S-W(1989)], in earlier works on the ergodicity of systems of elasctic hard balls, their masses were assumed to be identical. However, as it will be indicated in the first part of Section 6, it is very likely that the results of all these papers remain valid for systems of particles with different masses whenever the masses $m_1, \dots, m_N$ are independent over the rational field $\mathbb{Q}$.

Note 5. As to the basic results (and history) concerning the Boltzmann and Sinai hypotheses we refer to the recent survey [Sz (1996)]. The approach of the present work uses the following fundamental results: the Chernov-Sinai Fundamental Theorem (the Theorem on Local Ergodicity) for Semi-Dispersing Billiards [(S-Ch(1987)] in its form proved in [K-S-Sz(1990)] on one hand, and theorems of Simányi on the Uniform Avoiding of Balls and the appropriate generalization of his Connecting Path Formula on the other hand (see [Sim(1992)], where, in fact, the Boltzmann-Sinai Ergodic Hypothesis was already established for the case when $\nu \geq N$).

In the forthcoming part of the introduction the reader will be first reminded of the basic strategy for establishing global ergodicity of semi-dispersing billiards as it was suggested in our series of works with A. Krámli (see the references), and then we will point out the new elements of the method of the present work.

The basic notion in the theory of semi-dispersing billiards is that of the *sufficiency* of a phase point or, equivalently, of its orbit. The conceptual importance of sufficiency can be explained as follows (for a technical introduction and our prerequisites, see Section 2): In a suitably small neighbourhood of a (typical) phase point of a dispersing billiard the system is *hyperbolic*, i. e. its relevant Lyapunov exponents are not zero. For a semi-dispersing billiard the same property holds for sufficient points only! Physically speaking, a phase point is sufficient if its trajectory encounters in its history all possible degrees of freedom of the system. Then the theorem on local ergodicity for semi-dispersing billiards (see the main result of [S-Ch (1987)]) says that — under certain conditions — a suitably small neighbourhood of a sufficient phase point does belong to one ergodic component. As a consequence, if, for instance, almost every phase point of a semi-dispersing billiard is sufficient, then this property implies that the ergodic components are open and, therefore, their number is countable. To obtain, however, (global) ergodicity of the flow, a more stringent property of the subset of non-sufficient points is needed. For that purpose, in our work with A. Krámli [K-S-Sz (1989)] , the topological property of *slimness* (earlier misleadingly called 'residuality'), closely related to that of *topological codimension two*, was suggested and used in later works, too.

Definition 1.2. *We say that a subset of a smooth manifold M is* slim *if it can be covered by a countable union of codimension two (i. e. at least two), closed subsets of $\mu-$measure zero, where μ is a smooth measure on M.*

Remark. The phrase *dimension* in this paper always means the topological or Hurewicz dimension of a separable metric space, see for instance [H-W(1941)].

It has long been an accepted idea among experts that the right way of proving the K–property of hard ball systems is an induction on the number N of balls. According to our strategy initiated in [K-S-Sz (1991)] and explained in the introductions of [K-S-Sz (1992)] and of [S (1992)-I] , there are three major parts in such an induction, once a combinatorial property, called *richness* of the symbolic collision sequence of a trajectory, had been suitably defined.

(1) The "geometric–algebraic considerations" on the codimension of the manifolds describing the non–sufficient trajectory segments with a combinatorially rich symbolic collision structure (This step has no inductive character!);
(2) Proof of slimness (negligibility) of the set of phase points with a combinatorially non–rich collision sequence by using the K–property of less than N hard balls in the torus;
(3) Checking the Chernov–Sinai Ansatz — an important necessary condition for the proof of the fundamental theorem, see Condition (A) before Theorem 5 in [S-Ch (1987)] or Condition 3.1 in [K-S-Sz (1990)] . This step also uses the K–property of less than N balls and step (1) for the system of N hard balls in the torus.

As a consequence of the methods of [S(1992)], the main difficulty in our general case is to settle the geometric-algebraic part (1), and here we will also use some algebraic geometric tools. To do so (i) we have introduced *varying masses*, and can claim ergodicity for typical mass vectors, only; (ii) we will, in Section 3, *complexify the dynamics*. The complexification is required, on one hand, by the fact that our arguments in Section 4 use some algebraic geometric tools (see, for instance, the proof of Lemma 4.8) that assume the basic field to be algebraically closed. Another advantage of the complexification is that, in the inductive derivation of Key Lemma 4.1, one does not have to worry about the sheer existence of orbit segments with a prescribed symbolic collision sequence: they do exist, thanks to the algebraic closedness of the complex field.

The paper is organized as follows: Section 2 is devoted to prerequisites. Section 3 then describes the complexified dynamics, while in Section 4 we establish Key Lemma 4.1 — both for the complexified and the real dynamics — settling, in particular, part (1) of the strategy. Section 5 provides the demonstration of our Main Theorem, and, finally, Section 6 contains some comments and remarks.

2. Prerequisites

Semi-dispersing billiards and hard ball systems

Our approach is based on a simple observation:

Lemma 2.1. *The standard billiard ball flow with mass vector $\vec{m} \in \mathbb{R}^N_+$ is isomorphic to a semi-dispersing billiard.*

Because of the importance of the statement we sketch the proof.

Proof. Introduce new coordinates in the phase space as follows: For $(Q, V) \in \mathbf{M}$ let $\hat{q}_i = \sqrt{m_i} q_i, \;\; \hat{v}_i = \sqrt{m_i} v_i$, where thus $\hat{q}_i \in \sqrt{m_i}\ \mathbb{T}^\nu$ and $\sum_{i=1}^N \hat{v}_i^2 = 1$. Moreover, the vector $\hat{V} = (\hat{v}_1, \dots, \hat{v}_N)$ necessarily belongs to the hyperplane $\sum_{i=1}^N \sqrt{m_i} \hat{v}_i = 0$. Denote $\mathbb{T}_{\vec{m}} = \prod_{i=1}^N (\sqrt{m_i}\ \mathbb{T}^\nu)$ and

$$\widehat{C}_{i,j} = \left\{ \hat{Q} = (\hat{q}_1, \dots, \hat{q}_N) \in \mathbb{T}_{\vec{m}} : \;\; \left\| \frac{\hat{q}_i}{\sqrt{m_i}} - \frac{\hat{q}_j}{\sqrt{m_j}} \right\| < 2r \right\} \tag{2.2}$$

and $\tilde{\tilde{\hat{\mathbf{Q}}}} = \mathbb{T}_{\vec{m}} \setminus \cup_{1 \le i < j \le N} \hat{C}_{i,j}$. The equivalence relation $\hat{\Psi}_{\vec{m}}$ over $\mathbb{T}_{\vec{m}}$: $\hat{Q} \sim_{\hat{\Psi}_{\vec{m}}} \hat{Q}^*$ if and only if there exists an $a \in \mathbb{T}^\nu$ such that for every $i \in [1, N]$ $\;\hat{q}_i^* = \hat{q}_i + \sqrt{m_i} a$ allows us to define $\hat{\mathbf{Q}} = \tilde{\tilde{\hat{\mathbf{Q}}}} / \hat{\Psi}_{\vec{m}}$. Let $\hat{\mathbf{M}}_{\vec{m}} = \hat{\mathbf{Q}} \times \mathcal{S}^{d-1}$ be the unit tangent bundle of $\hat{\mathbf{Q}}$, and denote by $d\hat{\mu}$ the probability measure $const \cdot d\hat{Q} \cdot d\hat{V}$, where $d\hat{V}$ is the surface measure on the $(d-1)-$sphere $\mathcal{S}^{d-1}$, and $d\hat{Q}$ is the Lebesgue-measure on $\mathbb{T}_{\vec{m}}$. Now the standard billiard ball system $(\mathbf{M}, S^{\mathbb{R}}, \mu)_{\vec{m}}$ with mass vector $\vec{m}$ and the billiard system $(\hat{\mathbf{M}}_{\vec{m}}, \hat{S}^{\mathbb{R}}, \hat{\mu})$ are isomorphic.

Indeed, we can reduce the question to the case of one-dimensional particles because for both models the velocity components perpendicular to the normal of impact remain unchanged. The claimed isomorphy for one-dimensional particles, however, is well-known, and for its simple proof we can refer, for instance, to Section 4 of Chapter 5 in [C-F-S(1981)]. Thus the point is that in the isomorphic flow $\left(\hat{\mathbf{M}}_{\vec{m}}, \hat{S}^{\mathbb{R}}, \hat{\mu}\right)$ the velocity transformations at collisions become orthogonal reflections across the tangent hyperplane of the boundary of $\hat{\mathbf{Q}}$, see also (3.9) for the mentioned velocity transformation.

The fact that the billiard system $(\hat{\mathbf{M}}_{\vec{m}}, \hat{S}^{\mathbb{R}}, \hat{\mu})$ is semi-dispersing, is obvious, because the scattering bodies in $\hat{\mathbf{Q}}$ are cylinders built on ellipsoid bases (therefore this system is a cylindric billiard as introduced in [Sz(1993)]). $\square$

For convenience and brevity, we will throughout use the concepts and notations, related to semi-dispersing billiards and hard ball systems, of the papers [K-S-Sz(1990)] and [Sim(1992)-I-II], respectively, and will only point out where and how different masses play a role.

Remark 2.3 By slightly generalizing the notion of semi-dispersing billiards with allowing Riemannian metrics different from the usual ones (see (2.2) of [K-S-Sz(1990)]), we could have immediately identified the standard billiard ball system $(\mathbf{M}, S^{\mathbb{R}}, \mu)_{\vec{m}}$ with mass vector $\vec{m}$ with a (generalized) semi-dispersing billiard. (In the coordinates of the proof of Lemma 2.1, the Riemannian metric is

$$(d\rho)^2 = \sum_{i=1}^{N} \|d\hat{q}_i\|^2 + \|ds\|^2$$

where $\|ds\|^2$ is the square of the natural Riemannian metric on the unit sphere $\mathcal{S}^{d-1}$.) Then the standard billiard ball system $(\mathbf{M}, S^{\mathbb{R}}, \mu)_{\vec{m}}$ itself will be a (generalized) semi-dispersing billiard. The advantage is that — as it is easy to see — the results of [Ch-S(1987)] and of [K-S-Sz(1990)] remain valid for this class, too. For simplifying our exposition therefore, we will omit the change of coordinates of Lemma 2.1 and will be using the notion of semi-dispersing billiards in this slightly more general sense. In this sense the model of the Introduction is a semi-dispersing billiard.

An often used abbreviation is the shorthand $S^{[a,b]}x$ for the trajectory segment $\{S^t x : a \le t \le b\}$. The natural projections from $\mathbf{M}$ onto its factor spaces are denoted, as usual, by $\pi : \mathbf{M} \to \mathbf{Q}$ and $p : \mathbf{M} \to \mathcal{S}^{N\cdot\nu-\nu-1}$ or, sometimes, we simply write $\pi(x) = Q(x) = Q$ and $p(x) = V(x) = V$ for $x = (Q, V) \in \mathbf{M}$. Any $t \in [a, b]$ with $S^t x \in \partial M$ is called *a collision moment or collision time.*

As it is pointed out in previous works on billiards, the dynamics can only be defined for trajectories where the moments of collisions do not accumulate in any finite time interval (cf. Condition 2.1 of [K-S-Sz(1990)]). An important consequence of Theorem 5.3 of [V(1979)] is that — for semi-dispersing billiards! — there are *no trajectories at all with a finite accumulation point of collision moments* (see also [G(1981)] and [B-F-K(1995)]).

As a result, for an arbitrary non-singular orbit segment $S^{[a,b]}x$ of the standard billiard ball flow, there is a uniquely defined maximal sequence $a \le \tau_1 < \tau_2 < \cdots < \tau_m \le b :\ m \ge 0$ of collision times and a uniquely defined sequence $\sigma_1 < \sigma_2 < \cdots < \sigma_m$ of "colliding pairs", i. e. $\sigma_k = \{i_k, j_k\}$ whenever $Q(\tau_k) = \pi(S^{\tau_k}x) \in \partial\tilde{C}_{i_k,j_k}$. The sequence $\Sigma := \Sigma(S^{[a,b]}x) := (\sigma_1, \sigma_2, \ldots, \sigma_m)$ is called the *symbolic collision sequence* of the trajectory segment $S^{[a,b]}x$.

Definition 2.4. *We say that the symbolic collision sequence* $\Sigma = (\sigma_1, \dots, \sigma_m)$ *is connected if the collision graph of this sequence:*

$$\mathcal{G}_\Sigma := (\mathcal{V} = \{1, 2, \dots, N\}, \mathcal{E}_\Sigma := \{\{i_k, j_k\} : \text{ where } \sigma_k = \{i_k, j_k\}, \ 1 \le k \le m\})$$

is connected.

Definition 2.5. *We say that the symbolic collision sequence* $\Sigma = (\sigma_1, \dots, \sigma_m)$ *is* C*-rich, with* C *being a natural number, if it can be decomposed to at least* C *consecutive, disjoint collision sequences in such a way that each of them is connected.*

Neutral Subspaces, Advance and Sufficiency

Consider a *non–singular* trajectory segment $S^{[a,b]}x$. Suppose that a and b are *not moments of collision.* Before defining the neutral linear space of this trajectory segment, we note that the tangent space of the configuration space $\mathbf{Q}$ at interior points can be identified with the common linear space

$$\mathcal{Z} = \left\{ (w_1, w_2, \dots, w_N) \in (\mathbb{R}^\nu)^N : \sum_{i=1}^N m_i w_i = 0 \right\}. \tag{2.6}$$

Definition 2.7. *The neutral space* $\mathcal{N}_0(S^{[a,b]}x)$ *of the trajectory segment* $S^{[a,b]}x$ *at time zero* ($a < 0 < b$) *is defined by the following formula:*

$$\begin{aligned}\mathcal{N}_0(S^{[a,b]}x) = \big\{ W \in \mathcal{Z} : \exists(\delta > 0) \ s.\ t.\ \forall \alpha \in (-\delta, \delta) \\ p\left(S^a\left(Q(x) + \alpha W, V(x)\right)\right) = p(S^a x)\ \& \\ p\left(S^b\left(Q(x) + \alpha W, V(x)\right)\right) = p(S^b x)\big\}.\end{aligned}$$

It is known (see (3) in Section 3 of [S-Ch (1987)]) that $\mathcal{N}_0(S^{[a,b]}x)$ is a linear subspace of $\mathcal{Z}$ indeed, and $V(x) \in \mathcal{N}_0(S^{[a,b]}x)$. The neutral space $\mathcal{N}_t(S^{[a,b]}x)$ of the segment $S^{[a,b]}x$ at time $t \in [a, b]$ is defined as follows:

$$\mathcal{N}_t(S^{[a,b]}x) = \mathcal{N}_0\left(S^{[a-t,b-t]}(S^t x)\right). \tag{2.8}$$

It is clear that the neutral space $\mathcal{N}_t(S^{[a,b]}x)$ can be canonically identified with $\mathcal{N}_0(S^{[a,b]}x)$ by the usual identification of the tangent spaces of $\mathbf{Q}$ along the trajectory $S^{(-\infty,\infty)}x$ (see, for instance, Section 2 of [K-S-Sz (1990)]).

Our next definition is that of the *advance.* Consider a non-singular orbit segment $S^{[a,b]}x$ with symbolic collision sequence $\Sigma = (\sigma_1, \dots, \sigma_m)$ ($m \ge 1$) as at the beginning of the present section. For $x = (Q, V) \in \mathbf{M}$ and $W \in \mathcal{Z}$, $\|W\|$ sufficiently small, denote $T_W(Q, V) := (Q + W, V)$.

Definition 2.9. *For any* $1 \le k \le m$ *and* $t \in [a, b]$, *the advance*

$$\alpha(\sigma_k) : \ \mathcal{N}_t(S^{[a,b]}x) \to \mathbb{R}$$

is the unique linear extension of the linear functional defined in a sufficiently small neighbourhood of the origin of $\mathcal{N}_t(S^{[a,b]}x)$ *in the following way:*

$$\alpha(\sigma_k)(W) := \tau_k(x) - \tau_k(S^{-t} T_W S^t x).$$

It is now time to bring up the basic notion of *sufficiency* of a trajectory (segment). This is the utmost important necessary condition for the proof of the fundamental theorem for semi–dispersing billiards, see Condition (ii) of Theorem 3.6 and Definition 2.12 in [K-S-Sz (1990)] .

Definition 2.10.

(1) *The non–singular trajectory segment $S^{[a,b]}x$ (a and b are supposed not to be moments of collision) is said to be sufficient if and only if the dimension of $\mathcal{N}_t(S^{[a,b]}x)$ ($t \in [a,b]$) is minimal, i.e. dim $\mathcal{N}_t(S^{[a,b]}x) = 1$.*

(2) *The trajectory segment $S^{[a,b]}x$ containing exactly one singularity is said to be sufficient if and only if both branches of this trajectory segment are sufficient.*

For the notion of trajectory branches see, for example, the end of Section 2 in [S(1992)-I].

Definition 2.11. *The phase point $x \in \mathbf{M}$ with at most one singularity is said to be sufficient if and only if its whole trajectory $S^{(-\infty,\infty)}x$ is sufficient, which means, by definition, that some of its bounded segments $S^{[a,b]}x$ is sufficient.*

In the case of an orbit $S^{(-\infty,\infty)}x$ with exactly one singularity, sufficiency requires that both branches of $S^{(-\infty,\infty)}x$ be sufficient.

Connecting Path Formula for particles with different masses

The Connecting Path Formula, abbreviated as CPF, was discovered for particles with identical masses in [S (1992)-II] . Its goal was to give an explicit description (a linear coordinatization) of the neutral linear space $\mathcal{N}_0(S^{[-T,0]}x_0)$ in the language of the "advances" of the occuring collisions by using, as coefficients, linear expressions of the (pre-collision and post-collision) velocity differences of the colliding paticles. Since it relied upon the conservation of the momentum, it has been natural to expect that the CFP can be generalized for particles with different masses as well. The case is, indeed, this, and next we give this generalization for particles with different masses. Since its structure is the same as of the CPF for identical masses, our exposition follows closely that of [S (1992)-II] .

Consider a phase point $x_0 \in \mathbf{M}$ whose trajectory segment $S^{[-T,0]}x_0$ is not singular, $T > 0$. In the forthcoming discussion the phase point x_0 and the positive number T will be fixed. All the velocities $v_i(t) \in \mathbb{R}^\nu$ $i \in \{1, 2, \dots, N\}$, $-T \le t \le 0$ appearing in the considerations are velocities of certain balls at specified moments t and always with the starting phase point x_0. ($v_i(t)$ is the velocity of the i-th ball at time t.) We suppose that the moments 0 and $-T$ are not moments of collision. We label the balls by the natural numbers $1, 2, \dots, N$ (so the set $\{1, 2, \dots, N\}$ is always the vertex set of the collision graph) and we denote by $e_1, e_2, \dots, e_n$ the collisions of the trajectory segment $S^{[-T,0]}x_0$ (i.e. the edges of the collision graph) so that the time order of these collisions is just the opposite of the order given by the indices. Some more definitions and notations:

1. $\tau_i = \tau(e_i)$ denotes the time of the collision e_i, so $0 > \tau_1 > \tau_2 > \dots > \tau_n > -T$.
2. If $t \in \mathbb{R}$ is not a moment of collision ($-T \le t \le 0$), then

$$\Delta q_i(t) : \mathcal{N}_0(S^{[-T,0]}x_0) \to \mathbb{R}^\nu$$

is a linear mapping assigning to every element $W \in \mathcal{N}_0(S^{[-T,0]}x_0)$ the displacement of the i-th ball at time t, provided that the configuration displacement at time zero is given by W. Originally, this linear mapping is only defined for vectors

$W \in \mathcal{N}_0(S^{[-T,0]}x_0)$ close enough to the origin, but it can be uniquely extended to the whole space $\mathcal{N}_0(S^{[-T,0]}x_0)$ by preserving the linearity.

3. $\alpha(e_i)$ denotes the advance of the collision e_i, thus

$$\alpha(e_i):\ \mathcal{N}_0(S^{[-T,0]}x_0) \to \mathbb{R}$$

is a linear mapping $(i = 1, 2, \ldots, n)$.

4. The integers $1 = k(1) < k(2) < \cdots < k(l_0) \le n$ are defined by the requirement that for every j $(1 \le j \le l_0)$ the graph $\{e_1, e_2, \ldots, e_{k(j)}\}$ consists of $N - j$ connected components (on the vertex set $\{1, 2, \ldots, N\}$, as always!) while the graph $\{e_1, e_2, \ldots, e_{k(j)-1}\}$ consists of $N - j + 1$ connected components and, moreover, we require that the number of connected components of the whole graph $\{e_1, e_2, \ldots, e_n\}$ be equal to $N - l_0$. It is clear from this definition that the graph

$$\mathcal{T} = \{e_{k(1)}, e_{k(2)}, \ldots, e_{k(l_0)}\}$$

does not contain any loop, especially $l_0 \le N - 1$.

Here we make two remarks commenting the above notions.

Remark 2.12. We often do not indicate the variable $W \in \mathcal{N}_0(S^{[-T,0]}x_0)$ of the linear mappings $\Delta q_i(t)$ and $\alpha(e_i)$, for we will not be dealing with specific neutral tangent vectors W but, instead, we think of W as a typical (running) element of $\mathcal{N}_0(S^{[-T,0]}x_0)$ and $\Delta q_i(t)$, $\alpha(e_i)$ as linear mappings defined on $\mathcal{N}_0(S^{[-T,0]}x_0)$ in order to obtain an appropriate description of the neutral space $\mathcal{N}_0(S^{[-T,0]}x_0)$.

Remark 2.13. If $W \in \mathcal{N}_0(S^{[-T,0]}x_0)$ has the property $\Delta q_i(0)[W] = \lambda v_i(0)$ for some $\lambda \in \mathbb{R}$ and for all $i \in \{1, 2, \ldots, N\}$ (here $v_i(0)$ is the velocity of the i-th ball at time zero), then $\alpha(e_k)[W] = \lambda$ for all $k = 1, 2, \ldots, n$. This particular W corresponds to the direction of the flow. In the sequel we shall often refer to this remark.

Let us fix two *distinct* balls α, $\omega \in \{1, 2, \ldots, N\}$ that *are in the same connected component* of the collision graph $\mathcal{G}_n = \{e_1, e_2, \ldots, e_n\}$. The CPF expresses the relative displacement $\Delta q_\alpha(0) - \Delta q_\omega(0)$ in terms of the advances $\alpha(e_i)$ and the relative velocities occuring at these collisions e_i. In order to be able to formulate the CPF we need to define some graph-theoretic notions concerning the pair of vertices (α, ω).

Definition 2.14. *Since the graph $\mathcal{T} = \{e_{k(1)}, e_{k(2)}, \ldots, e_{k(l_0)}\}$ contains no loop and the vertices α, ω belong to the same connected component of $\mathcal{T}$, there is a unique path $\Pi(\alpha, \omega) = \{f_1, f_2, \ldots, f_h\}$ in the graph $\mathcal{T}$ connecting the vertices α and ω. The edges $f_i \in \mathcal{T}$ $(i = 1, 2, \ldots, h)$ are listed up successively along this path $\Pi(\alpha, \omega)$ starting from α and ending at ω. The vertices of the path $\Pi(\alpha, \omega)$ are denoted by $\alpha = B_0, B_1, B_2, \ldots, B_h = \omega$ indexed along this path going from α to ω, so the edge f_i connects the vertices B_{i-1} and B_i $(i = 1, 2, \ldots, h)$.*

When trying to compute $\Delta q_\alpha(0) - \Delta q_\omega(0)$ by using the advances $\alpha(e_i)$ and the relative velocities at these collisions, it turns out that not only the collisions f_i $(i = 1, 2, \ldots, h)$ make an impact on $\Delta q_\alpha(0) - \Delta q_\omega(0)$, but some other adjacent edges too. This motivates the following definition:

Definition 2.15. *Let $i \in \{1, 2, \dots, h-1\}$ be an integer. We define the set $\mathcal{A}_i$ of adjacent edges at the vertex B_i as follows:*

$$\mathcal{A}_i = \big\{e_j : j \in \{1, 2, \dots, n\} \;\&\; (\tau(e_j) - \tau(f_i)) \cdot (\tau(e_j) - \tau(f_{i+1})) < 0 \;\& \\ B_i \textit{ is a vertex of } e_j\big\}.$$

We adopt a similar definition to the sets $\mathcal{A}_0$, $\mathcal{A}_h$ of adjacent edges at the vertices B_0 and B_h, respectively:

Definition 2.16.

$$\mathcal{A}_0 = \big\{e_j : 1 \le j \le n \;\&\; \tau(e_j) > \tau(f_1) \;\&\; B_0 \textit{ is a vertex of } e_j\big\};$$

$$\mathcal{A}_h = \big\{e_j : 1 \le j \le n \;\&\; \tau(e_j) > \tau(f_h) \;\&\; B_h \textit{ is a vertex of } e_j\big\}.$$

We note that the sets $\mathcal{A}_0, \mathcal{A}_1, \dots, \mathcal{A}_h$ are not necessarily mutually disjoint.

Finally, we need to define the "contribution" of the collision e_j to $\Delta q_\alpha(0) - \Delta q_\omega(0)$ which is composed from the relative velocities just before and after the moment $\tau(e_j)$ of the collision e_j.

Definition 2.17. *For $i \in \{1, 2, \dots, h\}$ the "contribution" $\Gamma(f_i)$ of the edge $f_i \in \Pi(\alpha, \omega)$ is given by the formula*

$$\Gamma(f_i) = \begin{cases} v^-_{B_{i-1}}(\tau(f_i)) - v^-_{B_i}(\tau(f_i)), \\ \textit{if } \tau(f_{i-1}) < \tau(f_i) \;\&\; \tau(f_{i+1}) < \tau(f_i); \\ \\ v^+_{B_{i-1}}(\tau(f_i)) - v^+_{B_i}(\tau(f_i)), \\ \textit{if } \tau(f_{i-1}) > \tau(f_i) \;\&\; \tau(f_{i+1}) > \tau(f_i); \\ \\ \dfrac{1}{m_{B_{i-1}} + m_{B_i}} \big[m_{B_{i-1}}\big(v^-_{B_{i-1}}(\tau(f_i)) - v^-_{B_i}(\tau(f_i))\big) \\ +m_{B_i}\big(v^+_{B_{i-1}}(\tau(f_i)) - v^+_{B_i}(\tau(f_i))\big)\big], \\ \textit{if } \tau(f_{i+1}) < \tau(f_i) < \tau(f_{i-1}) \\ \\ \dfrac{1}{m_{B_{i-1}} + m_{B_i}} \big[m_{B_{i-1}}\big(v^+_{B_{i-1}}(\tau(f_i)) - v^+_{B_i}(\tau(f_i))\big) \\ +m_{B_i}\big(v^-_{B_{i-1}}(\tau(f_i)) - v^-_{B_i}(\tau(f_i))\big)\big], \\ \textit{if } \tau(f_{i-1}) < \tau(f_i) < \tau(f_{i+1}). \end{cases}$$

Here $v^-_{B_i}(\tau(f_i))$ denotes the velocity of the B_i-th particle just before the collision f_i (occuring at time $\tau(f_i)$) and, similarly, $v^+_{B_i}(\tau(f_i))$ is the velocity of the same particle just after the mentioned collision. We also note that, by convention, $\tau(f_0) = 0 > \tau(f_1)$ and $\tau(f_{h+1}) = 0 > \tau(f_h)$. We see that the time order plays an important role in this definition.

Definition 2.18. *For $i \in \{0,1,2,\dots,h\}$ the "contribution" $\Gamma_i(e_j)$ of an edge $e_j \in \mathcal{A}_i$ is defined as follows:*

$$\begin{aligned}\Gamma_i(e_j) = \;& sign\big(\tau(f_i)-\tau(f_{i+1})\big)\frac{m_C}{m_{B_i}+m_C}\cdot \\ & \cdot \Big[\big(v^+_{B_i}(\tau(e_j)) - v^+_C(\tau(e_j))\big) - \big(v^-_{B_i}(\tau(e_j)) - v^-_C(\tau(e_j))\big)\Big]\end{aligned}$$

where C is the vertex of e_j different from B_i.

Here again we adopt the convention of $\tau(f_0) = 0 > \tau(e_j)$ ($e_j \in \mathcal{A}_0$) and $\tau(f_{h+1}) = 0 > \tau(e_j)$ ($e_j \in \mathcal{A}_h$). We note that, by the definition of the set $\mathcal{A}_i$, exactly one of the two possibilities $\tau(f_{i+1}) < \tau(e_j) < \tau(f_i)$ and $\tau(f_i) < \tau(e_j) < \tau(f_{i+1})$ occurs. The subscript i of Γ is only needed because an edge $e_j \in \mathcal{A}_{i_1} \cap \mathcal{A}_{i_2}$ ($i_1 < i_2$) has two contributions at the vertices B_{i_1} and B_{i_2} which are just the endpoints of e_j.

We are now in the position of formulating the Connecting Path Formula:

Propostion 2.19. *Using all definitions and notations above, the following sum is an expression for $\Delta q_\alpha(0) - \Delta q_\omega(0)$ in terms of the advances and relative velocities of collisions:*

$$\Delta q_\alpha(0) - \Delta q_\omega(0) = \sum_{i=1}^{h} \alpha(f_i)\Gamma(f_i) + \sum_{i=0}^{h}\sum_{e_j \in \mathcal{A}_i} \alpha(e_j)\Gamma_i(e_j).$$

The proof of the propostion follows the proof of Simányi's CPF (Lemma 2.9 [S (1992)-II]) with the only difference that Lemma 2.8 of [S (1992)-II] is replaced here by the following

Lemma 2.20. *If e is a collision at time τ between the particles B and C, then*

$$v^+_B(\tau) - v^-_C(\tau) = \frac{m_C}{m_B+m_C}\big(v^+_B(\tau) - v^+_C(\tau)\big) + \frac{m_B}{m_B+m_C}\big(v^-_B(\tau) - v^-_C(\tau)\big)$$

and

$$v^+_B(\tau) - v^-_B(\tau) = \frac{m_C}{m_B+m_C}\Big[\big(v^+_B(\tau) - v^+_C(\tau)\big) - \big(v^-_B(\tau) - v^-_C(\tau)\big)\Big]$$

Proof. The lemma is an easy consequence of the conservation of momentum for the collision e. $\square$

Remark 2.21. In Section 3, we will complexify the system, and also allow complex masses, in particular. It is easy to see that the CPF of Proposition 2.19 and the whole discussion of this subsection will still hold for complex masses if we assume, in addition, that for every $e = \{B, C\}$, occuring in the symbolic collision sequence of $S^{[-T,0]}x_0$, $m_B + m_C \neq 0$.

3. The Complexified Billiard Map
Partial Linearity of the Dynamics

The aim of this section is to properly understand the algebraic relationship between the kinetic data of the billiard flow measured at different moments of time.

From now on we shall investigate orbit segments $S^{[0,T]}x_0$ $(T>0)$ of the standard billiard ball flow $\left(\tilde{\mathbf{M}}, \{S^t\}, \mu, \vec{m}\right)$, see the Introduction.

We note here that $\tilde{\mathbf{M}} = \tilde{\mathbf{Q}} \times \mathcal{E}$, where the configuration space $\tilde{\mathbf{Q}}$ was defined right after (1.1) and $\mathcal{E}$ is the velocity sphere

$$\mathcal{E} = \left\{ (v_1, \dots, v_N) \in \mathbb{R}^{\nu N} \middle| \sum_{i=1}^N m_i v_i = 0 \;\&\; \sum_{i=1}^N m_i \|v_i\|^2 = 1 \right\}$$

introduced in Section 1. Also note that in the geometric-algebraic considerations of the upcoming sections we do not use the equivalence relation Ψ of the introduction. Later on even the conditions $\sum_{i=1}^N m_i v_i = 0$ and $\sum_{i=1}^N m_i \|v_i\|^2 = 1$ will be dropped, cf. Remark 3.14.

The symbolic collision sequence of $S^{[0,T]}x_0$ is denoted by $\Sigma = \Sigma\left(S^{[0,T]}x_0\right) = (\sigma_1, \dots, \sigma_n)$. The symbol $v_i^k = v_i^k(t_k) \in \mathbb{R}^\nu$ shall denote the velocity $\dot{q}_i$ of the i-th ball right *after* the k-th collision $\sigma_k = \{i_k, j_k\}$ $(1 \le i_k < j_k \le N)$ occuring at time t_k, $k = 1, \dots, n$. Of course, there is no need to deal with the velocities right *before* collisions, since $v_i^-(t_k) = v_i^+(t_{k-1}) = v_i^{k-1}$. As usual, $q_i^k \in \mathbb{T}^\nu = \mathbb{R}^\nu / \mathbb{Z}^\nu$ denotes the position of the center of the i-th ball at the moment t_k of the k-th collision.

Let us now fix the symbolic collision sequence $\Sigma = (\sigma_1, \dots, \sigma_n)$ and explore the algebraic relationship between the data

$$\left\{ q_i^{k-1},\, v_i^{k-1} \middle| \, i = 1, \dots, N \right\}$$

and

$$\left\{ q_i^{k},\, v_i^{k} \middle| \, i = 1, \dots, N \right\},$$

$k = 1, \dots, n$. By definition, the data $\left\{ q_i^0,\, v_i^0 \middle| \, i = 1, \dots, N \right\}$ correspond to the initial (non-collision) phase point x_0. We also set $t_0 = 0$.

Since we would like to carry out arithmetic operations on these data, the periodic positions $q_i^k \in \mathbb{T}^\nu$ are not suitable for this purpose. Therefore, instead of studying the genuine orbit segments $S^{[0,T]}x_0$, we will deal with their Euclidean liftings.

Proposition 3.1. *Let $S^{[0,T]}x_0 = \{q_i(t),\, v_i(t) |\, 0 \le t \le T\}$ be an orbit segment as above, and assume that certain pre-images (Euclidean liftings) $\tilde{q}_i(0) = \tilde{q}_i^0 \in \mathbb{R}^\nu$ of the positions $q_i(0) \in \mathbb{T}^\nu = \mathbb{R}^\nu/\mathbb{Z}^\nu$ are given. Then there is a uniquely defined, continuous, Euclidean lifting $\{\tilde{q}_i(t) \in \mathbb{R}^\nu |\, 0 \le t \le T\}$ of the given orbit segment that is an extension of the initial lifting $\left\{\tilde{q}_i^0 \middle| \, i = 1, \dots, N\right\}$. Moreover, for every collision σ_k there exists a uniquely defined integer vector $a_k \in \mathbb{Z}^\nu$ – named the adjustment vector of σ_k – such that*

$$\left\| \tilde{q}_{i_k}^k - \tilde{q}_{j_k}^k - a_k \right\|^2 - 4r^2 = 0. \tag{3.2}$$

We shall call the orbit segment $\tilde{\gamma} = \left\{\tilde{q}_i^k,\, v_i^k \,|\, i=1,\dots,N;\; k=0,\dots,n\right\}$ *a lifted orbit segment with the system of adjustment vectors* $\mathcal{A} = (a_1,\dots,a_n) \in \mathbb{Z}^{n\nu}$.

The proof of this proposition is straightforward.

The next result establishes the already mentioned polynomial relationships between the kinetic data $\left\{\tilde{q}_i^{k-1},\, v_i^{k-1} \,|\, i=1,\dots,N\right\}$ and $\left\{\tilde{q}_i^k,\, v_i^k \,|\, i=1,\dots,N\right\}$, $k = 1,\dots,n$.

Proposition 3.3. *Using all notations and notions from above, one has the following polynomial relations between the kinetic data at* σ_{k-1} *and* σ_k, $k=1,\dots,n$. *In order to simplify the notations, we write down those equations as if* σ_k *was* $\{1,2\}$*:*

$$v_i^k = v_i^{k-1}, \quad i \notin \{1,2\}, \tag{3.4}$$

$$m_1 v_1^k + m_2 v_2^k = m_1 v_1^{k-1} + m_2 v_2^{k-1} \tag{3.5}$$

(conservation of the momentum),

$$v_1^k - v_2^k = v_1^{k-1} - v_2^{k-1} - \frac{1}{2r^2}\left\langle v_1^{k-1} - v_2^{k-1}; \tilde{q}_1^k - \tilde{q}_2^k - a_k\right\rangle \cdot \left(\tilde{q}_1^k - \tilde{q}_2^k - a_k\right) \tag{3.6}$$

(reflection of the relative velocity determined by the elastic collision),

$$\tilde{q}_i^k = \tilde{q}_i^{k-1} + \tau_k v_i^{k-1}, \quad i=1,\dots,N, \tag{3.7}$$

where the time slot $\tau_k = t_k - t_{k-1}$ $(t_0 := 0)$ *is determined by the quadratic equation*

$$\left\|\tilde{q}_1^{k-1} - \tilde{q}_2^{k-1} + \tau_k\left(v_1^{k-1} - v_2^{k-1}\right) - a_k\right\|^2 = 4r^2. \tag{3.8}$$

The proof of this proposition is also obvious.

Remark. Observe that the new velocities v_1^k and v_2^k can be computed from (3.5)-(3.6) as follows:

$$\begin{aligned} v_1^k &= v_1^{k-1} - \frac{m_2}{2r^2(m_1+m_2)}\left\langle v_1^{k-1} - v_2^{k-1}; \tilde{q}_1^k - \tilde{q}_2^k - a_k\right\rangle \cdot \left(\tilde{q}_1^k - \tilde{q}_2^k - a_k\right), \\ v_2^k &= v_2^{k-1} + \frac{m_1}{2r^2(m_1+m_2)}\left\langle v_1^{k-1} - v_2^{k-1}; \tilde{q}_1^k - \tilde{q}_2^k - a_k\right\rangle \cdot \left(\tilde{q}_1^k - \tilde{q}_2^k - a_k\right). \end{aligned} \tag{3.9}$$

We also note that the equations above even extend analytically to the case when one of the two masses, say m_1, is equal to zero:

$$\begin{aligned} v_1^k &= v_1^{k-1} - \frac{1}{2r^2}\left\langle v_1^{k-1} - v_2^{k-1}; \tilde{q}_1^k - \tilde{q}_2^k - a_k\right\rangle \cdot \left(\tilde{q}_1^k - \tilde{q}_2^k - a_k\right), \\ v_2^k &= v_2^{k-1}. \end{aligned} \tag{3.10}$$

We also note that the symbol $\langle\, .\, ;\, .\,\rangle$ denotes the Euclidean inner product of ν-dimensional real vectors and $\|\,.\,\|$ is the corresponding norm.

Given the pair $(\Sigma, \mathcal{A}) = (\sigma_1, \dots, \sigma_n; a_1, \dots, a_n)$, the equations (3.4)-(3.8) make it possible to iteratively compute the kinetic data $\{\tilde{q}_i^k, v_i^k | i = 1, \dots, N\}$ by using the preceding data $\{\tilde{q}_i^{k-1}, v_i^{k-1} | i = 1, \dots, N\}$. Throughout these computations one only uses the field operations and the square root. (The latter one is merely used when computing τ_k as the root of the quadratic equation (3.8). Obviously, since – at the moment – we are dealing with genuine, real orbit segments, in this dynamically realistic situation the equations (3.8) have two distinct, positive real roots, and τ_k is the smaller one.) Therefore, the data $\{\tilde{q}_i^k, v_i^k | i = 1, \dots, N\}$ can be expressed as certain algebraic functions of the initial data $\{\tilde{q}_i^0, v_i^0 | i = 1, \dots, N\}$, and the arising algebraic functions merely contain the field operations, the square root and, finally, the radius r of the balls and the masses m_i as constants.

Since these algebraic functions make full sense over the complex field $\mathbb{C}$ and, after all, our proof of the Theorem requires the complexification, we are now going to complexify the whole system by considering the kinetic variables and the masses as complex ones and retaining the polynomial equations (3.4)-(3.8). However, due to the ambiguity of selecting a root of (3.8) out of the two, it proves to be important to explore first the algebraic frame of the relations (3.4)-(3.8) connecting the studied variables. This is what we do now.

The Field Extension Associated with the Pair $(\Sigma, \mathcal{A})$

We are going to define the commutative function field $\mathbb{K} = \mathbb{K}(\Sigma; \mathcal{A})$ generated by all functions

$$\left\{(\tilde{q}_i^k)_j,\ (v_i^k)_j \middle| i = 1, \dots, N;\ k = 0, \dots, n;\ j = 1, \dots, \nu\right\}$$

of the lifted orbit segments corresponding to the given parameters

$$(\Sigma, \mathcal{A}) = (\sigma_1, \dots, \sigma_n; a_1, \dots, a_n)$$

in such a way that the field $\mathbb{K}(\Sigma; \mathcal{A})$ incorporates all algebraic relations among these variables that are consequences of the equations (3.4)-(3.8). (Here the subscript j denotes the j-th component of a ν-vector.) In our setup the ground field of allowed constants (coefficients) is, by definition, the complex field $\mathbb{C}$. The precise definition of $\mathbb{K}_n = \mathbb{K}(\Sigma; \mathcal{A})$ is

Definition 3.11. *For $n = 0$ the field $\mathbb{K}_0 = \mathbb{K}(\emptyset; \emptyset)$ is the transcendental extension $\mathbb{C}(\mathcal{B})$ of the coefficient field $\mathbb{C}$ by the algebraically independent formal variables*

$$\mathcal{B} = \left\{(\tilde{q}_i^0)_j,\ (v_i^0)_j,\ m_i \middle| i = 1, \dots, N;\ j = 1, \dots, \nu\right\}.$$

Suppose now that $n > 0$ and the commutative field $\mathbb{K}_{n-1} = \mathbb{K}(\Sigma'; \mathcal{A}')$ has already been defined, where $\Sigma' = (\sigma_1, \dots, \sigma_{n-1})$, $\mathcal{A}' = (a_1, \dots, a_{n-1})$. Then we consider the quadratic equation $b_n\tau_n^2 + c_n\tau_n + d_n = 0$ in (3.8) with $k = n$ as a polynomial equation defining a new field element τ_n to be adjoined to the field $\mathbb{K}_{n-1} = \mathbb{K}(\Sigma'; \mathcal{A}')$. (Recall that the coefficients b_n, c_n, d_n come from the field $\mathbb{K}_{n-1}$.)

There are two possibilities:

(i) The quadratic polynomial $b_n x^2 + c_n x + d_n$ *is reducible over the field* $\mathbb{K}_{n-1}$*. Then we take* $\mathbb{K}_n = \mathbb{K}_{n-1}$*.*

(ii) The polynomial $b_n x^2 + c_n x + d_n$ *is irreducible over the field* $\mathbb{K}_{n-1}$*. Then we define* $\mathbb{K}_n = \mathbb{K}(\Sigma; \mathcal{A})$ *as the extension of* $\mathbb{K}_{n-1} = \mathbb{K}(\Sigma'; \mathcal{A}')$ *by the root* τ_n *of this irreducible polynomial.*

Remark 3.12. The importance of the field $\mathbb{K}_n$ is underscored by the fact that this field encodes all algebraic relations among the kinetic data that follow from the polynomial equations (3.4)-(3.8). Furthermore, Proposition 3.3 gives us an iterative computation rule for successively obtaining the kinetic variables

$$\left\{\tilde{q}_i^k,\, v_i^k \,|\, i = 1, \dots, N\right\},$$

$k = 0, \dots, n$. The field $\mathbb{K}_n = \mathbb{K}(\Sigma, \mathcal{A})$ is, after all, the algebraic frame of such computations. We note that – everywhere in this paper – the symbols $\langle\, .\, ;\, .\,\rangle$ and $\|\, .\, \|^2$ do not refer to an Hermitian inner product but, rather, $\langle x; y\rangle = \sum_{j=1}^{\nu} x_j y_j$ and $\|x\|^2 = \sum_{j=1}^{\nu} x_j^2$, so that these expressions retain their polynomial form.

Remark 3.13. Later on it will be necessary to express each kinetic variable $\left(\tilde{q}_i^k\right)_j$ and $\left(v_i^k\right)_j$ as an algebraic function of the initial variables of $\mathcal{B}$. This raises, however, an important question: Which one of the two roots τ_k of (3.8) should be considered during these computations? This is no problem if the polynomial $b_k x^2 + c_k x + d_k$, defining τ_k, is irreducible over the field $\mathbb{K}_{k-1}$, because the two roots of this polynomial are algebraically equivalent over $\mathbb{K}_{k-1}$, and any of them can be chosen as τ_k. However, if the polynomial $b_k x^2 + c_k x + d_k$ is reducible over the field $\mathbb{K}_{k-1}$, then it is necessary to make a decision and assign to one of the two roots (both in the field $\mathbb{K}_{k-1}$ now) the role of τ_k. This is what we do. The result is the field $\mathbb{K}_n = \mathbb{K}(\Sigma; \mathcal{A})$ endowed with a distinguished n-tuple $(\tau_1, \dots, \tau_n)$ of its elements. The field $\mathbb{K}_n$ with the distinguished n-tuple $(\tau_1, \dots, \tau_n) = \vec{\tau}$ is denoted by $\mathbb{K}(\Sigma; \mathcal{A}; \vec{\tau})$. Then the algebraic object $\mathbb{K}(\Sigma; \mathcal{A}; \vec{\tau})$ completely controls the whole process of the iterative computation of the kinetic variables.

Note that in the genuine, real case the two roots of $b_k x^2 + c_k x + d_k$ are distinct, positive real numbers, and – by the nature of the billiard dynamics – the chosen value is always the smaller one.

Remark 3.14. The reader may wonder why we did not postulate the algebraic dependencies $\sum_{i=1}^{N} m_i v_i^k = 0$, $\sum_{i=1}^{N} m_i \|v_i^k\|^2 = 1$? The answer is the following: The definition of the neutral linear space together with its characterization via the Connecting Path Formula (CPF, see Section 2 of this paper or Lemma 2.9 of [Sim(1992)-II]), i. e. the partial linearity of the billiard flow, are invariant under

(1) all uniform velocity translations (adding the same vector to all velocities) and

(2) time rescalings.

Remark 3.15. All fields occuring in this paper are only defined up to an isomorphism over the coefficient field $\mathbb{C}$. Therefore, the statements like "the field $\mathbb{K}_2$ is an extension of $\mathbb{K}_1$" should be understood as follows: $\mathbb{K}_1$ is a subfield of $\mathbb{K}_2$ after the natural identification of the generating variables bearing the same name. For the necessary notions, properties, and results from the theory of field extensions and Galois theory, the reader is kindly recommended to look up, for instance, the book by I. Stewart, [St(1973)].

Remark 3.16. As said before (see Definition 3.11), the collection $\mathcal{B}$ of elements of the field $\mathbb{K}(\Sigma;\mathcal{A};\vec{\tau})$ is a base of transcendence in that field over the subfield $\mathbb{C}$, and the degree of the extension $\mathbb{K}(\Sigma;\mathcal{A};\vec{\tau}) : \mathbb{C}(\mathcal{B}) = \mathbb{K}_n : \mathbb{K}_0$ is a power of two. Moreover, each of the following sets is a generator for the field $\mathbb{K}_n$:

(a) $\mathbb{C}\cup\mathcal{B}\cup\{\tau_1,\dots,\tau_n\}$;

(b) $\mathbb{C}\cup\{m_1,\dots,m_N\}\cup\{(\tilde{q}_i^n)_j,\,(v_i^n)_j\,|\,i=1,\dots,N;\ j=1,\dots,\nu\}\cup\{\tau_1,\dots,\tau_n\}$;

(c) $\mathbb{C}\cup\{m_1,\dots,m_N\}\cup\big\{(\tilde{q}_i^k)_j,\,(v_i^k)_j\,|\,i=1,\dots,N;\ j=1,\dots,\nu;\ k=0,\dots,n\big\}$.

Remark 3.17. The above procedure of constructing field extensions is closely related to the classical theory of geometric constructions by a ruler and compass, and this is not for surprise: the billiard trajectory is constructed by intersecting a straight line with a sphere and then reflecting it across the tangent hyperplane of the sphere. This is a sort of classical geometric construction in ν dimensions. For the details see, for instance, §57 and §60 of [VDW (1955)], or Chapter 5 of [St(1973)].

Let us fix the pair $(\Sigma;\mathcal{A})$ and the n-tuple $\vec{\tau}=(\tau_1,\dots,\tau_n)$ of elements of the field $\mathbb{K}(\Sigma;\mathcal{A})$. We are now defining a $(2\nu+1)N$-dimensional complex analytic manifold $\tilde{\Omega}=\tilde{\Omega}\,(\Sigma;\mathcal{A};\vec{\tau})$ and certain holomorphic functions $\tilde{q}_i^k,\,v_i^k:\tilde{\Omega}\to\mathbb{C}^\nu$, $m_i:\tilde{\Omega}\to\mathbb{C}$, $\tau_k:\tilde{\Omega}\to\mathbb{C}$, $i=1,\dots,N$, $k=0,\dots,n$, (τ_0 is not defined). First we define a domain $D=D\,(\Sigma;\mathcal{A};\vec{\tau})\subset\mathbb{C}^{(2\nu+1)N}$ as the set of all complex $(2\nu+1)N$-tuples

$$\big\{(\tilde{q}_i^0)_j,\,(v_i^0)_j,\,m_i|\,i=1,\dots,N;\ j=1,\dots,\nu\big\}$$

for which

(a) the leading coefficient $\left\|v_{i_k}^{k-1}-v_{j_k}^{k-1}\right\|^2$ of the quadratic equation (3.8) is never zero, $k=1,\dots,n$,

(b) $m_{i_k}+m_{j_k}\neq 0$ for $k=1,\dots,n$,

provided that we carry out the iterative computation of the kinetic variables

$$\big\{(\tilde{q}_i^k)_j,\,(v_i^k)_j|\,i=1,\dots,N;\ j=1,\dots,\nu;\ k=0,\dots,n\big\}$$

as governed by the polynomial equations (3.4)-(3.8) and by the fixed selection of elements $\vec{\tau}$.

Remark 3.17/a. Note that we require the validity of (a)-(b) for any branch of taking the square roots when computing the τ_k's with irreducible defining polynomials (3.8).

Plainly, the complement set $\mathbb{C}^{(2\nu+1)N}\setminus D$ is a proper, closed, algebraic subset of $\mathbb{C}^{(2\nu+1)N}$, especially a finite union of complex analytic submanifolds with codimension at least one, so the open set $D\subset\mathbb{C}^{(2\nu+1)N}$ is *connected and dense.*

Definition 3.18. *We define $\tilde{\Omega}\,(\Sigma;\mathcal{A};\vec{\tau})$ as the set of all complex $(2\nu(n+1)N+N+n)$-tuples*

$$\omega=\Big(\left(\tilde{q}_i^k\right)_j,\,(v_i^k)_j,\,m_i,\,\tau_l\Big|\;i=1,\dots,N;\;k=0,\dots,n;\;j=1,\dots,\nu;\;l=1,\dots,n\Big)$$

for which these coordinates are interrelated by the equations (3.4)–(3.8) and the vector

$$\vec{x}(\omega)=\Big(\left(\tilde{q}_i^0\right)_j,\,(v_i^0)_j,\,m_i\Big|\,i=1,\dots,N;\;j=1,\dots,\nu\Big)$$

of initial data belongs to the set $D\,(\Sigma;\mathcal{A};\vec{\tau})$ defined above.

It is clear now that $\tilde{q}_i^k,\,v_i^k\,:\,\tilde{\Omega}\to\mathbb{C}^\nu$, $m_i\,:\,\tilde{\Omega}\to\mathbb{C}$, and $\tau_k\,:\,\tilde{\Omega}\to\mathbb{C}$ $(i=1,\dots,N,\;k=0,\dots,n)$ are holomorphic functions on $\tilde{\Omega}=\tilde{\Omega}\,(\Sigma;\mathcal{A};\vec{\tau})$. The complex analytic manifold $\tilde{\Omega}$ endowed with the above holomorphic functions can justifiably considered as the complexification of the $(\Sigma;\mathcal{A};\vec{\tau})$-iterated billiard map.

Remark. It is very likely that the manifold $\tilde{\Omega}$ is connected. Nevertheless, in our proof we do not need this connectedness, and are not going to pursue the goal of proving it.

The Complex Neutral Space $\mathcal{N}(\omega)$

Fix a base point $\omega\in\tilde{\Omega}\,(\Sigma;\mathcal{A};\vec{\tau})$. The tangent space $\mathcal{T}_\omega\tilde{\Omega}$ of $\tilde{\Omega}$ at ω consists of all complex vectors

$$x=\left(\delta\tilde{q}_i^0,\,\delta v_i^0,\,\delta m_i\right|i=1,\dots,N\big)\in\mathbb{C}^{(2\nu+1)N},$$

thus $\mathcal{T}_\omega\tilde{\Omega}$ can be naturally identified with the complex vector space $\mathbb{C}^{(2\nu+1)N}$. Set

$$\mathcal{N}(\omega)=\Big\{x=\left(\delta\tilde{q}_i^0,\,\delta v_i^0,\,\delta m_i\right|i=1,\dots,N\big)\in\mathcal{T}_\omega\tilde{\Omega}\Big|\;D_x(v_i^k)=D_x(m_i)=0\text{ for }i=1,\dots,N;\;k=0,\dots,n\Big\},\tag{3.19}$$

where $D_x(\,.\,)$ denotes the directional differentiation in the direction of x. Clearly, the complex linear subspace $\mathcal{N}(\omega)$ of $\mathbb{C}^{(2\nu+1)N}$ is the proper complex analogue of the neutral space of a genuine, real orbit segment.

Remark. Since δv_i^0 and δm_i must be equal to zero for a neutral vector $x\in\mathcal{N}(\omega)$, we shall simply write $x=(\delta\tilde{q}_1^0,\dots,\delta\tilde{q}_N^0)$ instead of indicating the zero entries.

The proof of the following proposition is completely analogous to that one for the real case, and therefore we omit it.

Proposition 3.20. *For every tangent vector $x \in \mathcal{N}(\omega)$ $(\omega \in \tilde{\Omega}(\Sigma;\mathcal{A};\vec{\tau}))$ and $1 \le k \le n$ we have*

$$D_x\left(\tilde{q}_{i_k}^{k-1} - \tilde{q}_{j_k}^{k-1}\right) = (\alpha_k(x) - \alpha_{k-1}(x)) \cdot \left(v_{i_k}^{k-1} - v_{j_k}^{k-1}\right), \tag{3.21}$$

$$D_x(t_k) = -\alpha_k(x), \tag{3.22}$$

and

$$D_x(\tilde{\tau}_k) = D_x(t_k - t_{k-1}) = \alpha_{k-1}(x) - \alpha_k(x), \tag{3.23}$$

where, by definition, $t_0 = 0 = \alpha_0(x)$. The functions $\alpha_k : \mathcal{N}(\omega) \to \mathbb{C}$ $(k = 1, \dots, n)$ are linear functionals. The name of the functional α_k is the advance of the k-th collision $\sigma_k = \{i_k, j_k\}$, see also Section 2.

This result pregnantly shows that, indeed, the vector space $\mathcal{N}(\omega)$ describes the linearity of the $(\Sigma;\mathcal{A};\vec{\tau})$-iterated billiard map.

Remark 3.24. It is a matter of simple computation to convince ourselves that all assertions of Section 2 of the present paper and [Sim(1992)-II] pertaining to the neutral space $\mathcal{N}(\omega)$ remain valid for the complexified dynamics. Here we only point out the two most important statements out of those results:

(1) The vector space $\mathcal{N}(\omega)$ measures the ambiguity in determining the orbit segment ω purely by its velocity history $\left\{v_i^k(\omega) : i = 1, \dots, N;\ k = 0, \dots, n\right\}$ and the masses $\vec{m}$. In other words, this means that if locally in the phase space $\tilde{\Omega}$ (in a small open set) two phase points ω_1 and ω_2 have the same velocities ($v_i^k(\omega_1) = v_i^k(\omega_2)$, $i = 1, \dots, N$; $k = 0, \dots, n$) and masses, then the initial data of ω_1 and ω_2 can only differ by a spatial translation by a vector from the space $\mathcal{N}(\omega_1) = \mathcal{N}(\omega_2)$, and this statement is obviously reversible.

(2) The Connecting Path Formula (CPF, Lemma 2.9 of [Sim(1992)-II], or Proposition 2.19 in this paper) is applicable, and it enables us to compute the neutral space $\mathcal{N}(\omega)$ via solving a homogeneous system of linear equations (now over the field $\mathbb{C}$)

$$\sum_{k=1}^{n} \alpha_k \Gamma_{ik} = 0,\ i = 1, \dots, n + P_\Sigma - N, \tag{3.25}$$

where each equation in (3.25) is a ν-dimensional complex vector equation, the coefficients $\Gamma_{ik} = \Gamma_{ik}(\omega) \in \mathbb{C}^\nu$ are certain rational linear combinations of relative velocities $v_i^k(\omega) - v_j^k(\omega)$ and, finally, P_Σ denotes the number of connected components of the collision graph of Σ.

Remark 3.26. As a straightforward consequence of the equation (3.6), we obtain that $v_i^{k-1} = v_i^k$ for $i = 1, \dots, N$ (i. e. the k-th reflection does not change the compound velocity) if and only if

$$\left\langle v_{i_k}^{k-1} - v_{j_k}^{k-1};\ \tilde{q}_{i_k}^k - \tilde{q}_{j_k}^k - a_k \right\rangle = 0,$$

and this is just the case of a tangential collision. Easy calculation shows that the mentioned tangentiality occurs if and only if the two roots τ_k of (3.8) coincide.

It is obvious that $\dim_{\mathbb{C}} \mathcal{N}(\omega)$ is at least $\nu + 1$ (as long as not all velocities are the same), because the flow direction and the uniform spatial translations are necessarily contained by $\mathcal{N}(\omega)$.

Definition 3.27 (Definition of sufficiency). *The orbit segment $\omega \in \tilde{\Omega}(\Sigma; \mathcal{A}; \vec{\tau})$ is said to be sufficient if and only if $dim_{\mathbb{C}}\mathcal{N}(\omega) = \nu + 1$.*

4. The Notion of Richness
Richness Implies Sufficiency
(The Case $\nu \geq 3$)

The result of this section is

Key Lemma 4.1. *Let $\nu \geq 3$. There exists a positive number $C(N)$ (depending merely on the number of balls $N \geq 2$) with the following property: If a symbolic sequence $\Sigma = (\sigma_1, \dots, \sigma_n)$ is $C(N)$-rich and the N-tuple of masses $\vec{m} = (m_1, \dots, m_N)$ does not belong to some zero measure subset of $\mathbb{C}^N$ (actually, does not belong to a countable union of proper, complex analytic submanifolds of $\mathbb{C}^N$), then $dim_{\mathbb{C}}\mathcal{N}(\omega) = \nu + 1$ for every $\omega \in \tilde{\Omega}(\Sigma, \mathcal{A}, \vec{\tau}, \vec{m})$ except some closed, algebraic subset E of $\tilde{\Omega}$ with complex codimension at least two. The real version of this result is also valid. (The manifold $\tilde{\Omega}(\Sigma, \mathcal{A}, \vec{\tau}, \vec{m})$ is the submanifold of $\tilde{\Omega}(\Sigma; \mathcal{A}; \vec{\tau})$ – whose definition can be found in 3.18 – corresponding to the selection of masses $\vec{m} = (m_1, \dots, m_N)$. As for the definition of a C-rich symbolic sequence, see 2.5.)*

As mentioned in the Introduction, the aim of Sections 3 and 4, in particular that of Key Lemma 4.1, is to settle part (1) of the strategy detailed there. The methods are mainly algebraic, and for an easier separation of our dynamical vs. algebraic arguments, we are now going to describe *what are the only properties of the dynamics (and nothing more) that we are going to use.*

(1) The dynamical properties of an orbit segment $S^{[0,T]}x_0$ with a symbolic collision sequence Σ are completely characterized by the equations (3.4-8) if we, in addition, specify τ_k as the smaller of the two, necessarily positive, roots of the equation (3.8). In the inductive steps of Lemmas 4.11 and 4.19 we also use the remark to Proposition 3.3 saying that the collision laws (3.9), themselves consequences of (3.5-6), extend analytically to the case of zero mass, cf. (3.10).

(2) Since our goal is to show that orbits with a rich collision structure are typically sufficient or, equivalently, their neutral subspaces have typically minimal possible dimension, we also need an algebraic characterization of the neutral subspace. This is attained through the CPF (see Proposition 2.19), and our arguments also use some simple inference about the CPF and the neutral subspace from [S (1992)-II] .

(3) As we have seen in Section 3, the algebraic formalisms of points (1) and (2) can be complexified, and our forthcoming discussion is based upon this complexification.

Remark. In the course of proving the key lemma we will see that the thresholds $C(N)$ $(N \geq 2)$ can actually be defined by the recursion $C(2) = 1$, $C(N) = \frac{N}{2} \cdot \max\{C(N-1), 3\}$ $(N \geq 3)$, that is, $C(2) = 1$, $C(N) = \frac{3N!}{2^{N-1}}$ for $N \geq 3$. We note that, as it has been proved in [K-S-Sz(1991)], for $N = 3$ the value $C(3) = 2$ already does the job.

Proof of Key Lemma 4.1. All the rest of this section is devoted to the inductive prof of the key lemma, and it will be split into several lemmas.

First of all, we prove the original, complex version of the key lemma by an induction on the number N. Once it is completed, we will make the necessary observations and remarks (Lemma 4.22) by showing that the validity of the key lemma carries over to the real case.

Lemma 4.2. *Consider the open, connected and dense subset D of $\mathbb{C}^{(2\nu+1)N}$ formed by all possible complex $(2\nu+1)N$-tuples*

$$\vec{x} = \vec{x}(\omega) = \left(\left(\tilde{q}_i^0(\omega)\right)_j, \left(v_i^0(\omega)\right)_j, m_i(\omega) \middle| \, i = 1, \dots, N; \; j = 1, \dots, \nu \right), \qquad \left(\omega \in \tilde{\Omega}\right)$$

of initial data of the complex $(\Sigma; \mathcal{A}; \vec{\tau})$-dynamics, see the definition of

$$\tilde{\Omega} = \tilde{\Omega}\left(\Sigma; \mathcal{A}; \vec{\tau}\right)$$

in the previous section. We claim that there are finitely many complex polynomials $P_1(\vec{x}), \dots, P_s(\vec{x})$ ($\vec{x} \in \mathbb{C}^{(2\nu+1)N}$) such that for every $\vec{x} \in D$ one has

$$\begin{aligned} &P_1(\vec{x}) = \cdots = P_s(\vec{x}) = 0 \Longleftrightarrow \\ \exists\, \omega \in \tilde{\Omega}\left(\Sigma; \mathcal{A}; \vec{\tau}\right) \;\; &\text{such that } \vec{x} = \vec{x}(\omega) \,\&\, \dim_{\mathbb{C}} \mathcal{N}(\omega) > \nu + 1. \end{aligned} \tag{4.3}$$

Remark. The right-hand-side in the above equivalence precisely says that *some* branch of the multiple-valued $(\Sigma; \mathcal{A}; \vec{\tau})$-dynamics with initial data $\vec{x}$ is not sufficient, see also Remark 3.17/a.

Proof. This proof is a typical application of the Connecting Path Formula (see Proposition 2.19 here or Lemma 2.9 of [Sim(1992)-II]) and Proposition 3.4 of the same article. Namely, Proposition 3.4 of [Sim(1992)-II] asserts that

$$\dim_{\mathbb{C}} \mathcal{N}(\omega) = \nu \cdot P_\Sigma + \dim_{\mathbb{C}} \{\alpha_1, \dots, \alpha_n\}, \tag{4.4}$$

where, as said before, P_Σ denotes the number of connected components of the collision graph of Σ, and $\{\alpha_1, \dots, \alpha_n\}$ is a shorthand for the complex linear space of all possible n-tuples $(\alpha_1(x), \dots, \alpha_n(x))$ of advances, $x \in \mathcal{N}(\omega)$.

Remark 4.5. It is worth noting here that our present formula (4.4) differs from its counterpart in Proposition 3.4 of [Sim(1992)-II] by an additional term ν. This is, however, due to the fact that in the present approach we no longer have the reduction equation $\sum_{i=1}^N \delta\tilde{q}_i = 0$. It follows easily from (4.4) that the sufficiency is equivalent to

$$\dim_{\mathbb{C}} \{\alpha_1, \dots, \alpha_n\} = 1,$$

i. e. the sufficiency means that all advances are equal to the same functional. Note that $\dim_{\mathbb{C}} \{\alpha_1, \dots, \alpha_n\} = 1$ obviously implies $P_\Sigma = 1$.

A simple, but important consequence of the Connecting Path Formula is that the linear space $\{\alpha_1, \dots, \alpha_n\}$ is the solution set of a homogeneous system of linear equations (3.25).

Remark 4.6. Note that, when applying the mentioned lemma, the left-hand-side of the Connecting Path Formula has to be written as the relative velocity of the colliding balls multiplied by the advance of that collision. It follows from the exposition of [Sim(1992)-II] that the equations of this sort (arising from all CPF's) are the only constraints on the advance functionals. Also note that the number $n+P_\Sigma-N$ of equations in (3.25) is equal to the number of all collisions $\sigma_k = \{i_k, j_k\}$ for which i_k and j_k are in the same connected component of the collision graph of $\{\sigma_1, \sigma_2, \dots, \sigma_{k-1}\}$.

The determinant $D(M)$ of every square-shaped submatrix M of the coefficient matrix of (3.25) is a homogeneous velocity polynomial and, therefore, $D(M)$ is a holomorphic function on $\tilde{\Omega}$. Thus, there are finitely many velocity polynomials (actually, determinants) $R_1(\omega), \dots, R_s(\omega)$ such that their simultaneous vanishing is equivalent to the non-sufficiency of the orbit segment ω. By using the equations (3.4)-(3.8) (which define the complex dynamics recursively), each of these velocity polynomials $R_i(\omega)$ can be – essentially uniquely – written as an algebraic function $f_i(\vec{x}(\omega)) = f_i(\vec{x})$ of the initial data $\vec{x}$, and these algebraic functions $f_i(\vec{x})$ only contain (finitely many) field operations and the square root. By using the canonical method of the successive elimination of the square roots from the equation $f_i(\vec{x}) = 0$, one obtains a complex polynomial $P_i(\vec{x})$ such that

$$P_i(\vec{x}) = 0 \iff f_i(\vec{x}) = 0.$$

We emphasize here that we understand the relation $f_i(\vec{x}) = 0$ in such a way that $f_i(\vec{x})$ becomes zero on *some* branch of evaluating the multiple-valued algebraic function $f_i(\vec{x})$, cf. Remark 3.17/a.

Remark 4.7. The process of eliminating the square roots from an equation $f(\vec{x}) = 0$ (of the above type), however, requires a bit of clarification. The algebraic function f has a natural representing element $\hat{f}$ in the field $\mathbb{K}_n$, see Definition 3.11. Suppose that $\hat{f} \in \mathbb{K}_k$ but $\hat{f} \notin \mathbb{K}_{k-1}$. Then the field element $\hat{f} \in \mathbb{K}_k$ can be written uniquely in the form $\hat{f} = a + b\sqrt{d}$, where $a, b, d \in \mathbb{K}_{k-1}$, $b \neq 0$, $\sqrt{d} \notin \mathbb{K}_{k-1}$, and

$$d = \left\langle v_{i_k}^{k-1} - v_{j_k}^{k-1};\ \tilde{q}_{i_k}^{k-1} - \tilde{q}_{j_k}^{k-1} - a_k \right\rangle^2 - \\ \left\| v_{i_k}^{k-1} - v_{j_k}^{k-1} \right\|^2 \cdot \left(\left\| \tilde{q}_{i_k}^{k-1} - \tilde{q}_{j_k}^{k-1} - a_k \right\|^2 - 4r^2 \right)$$

is the discriminant of the quadratic equation (3.8) with the unknown τ_k.

In the first step of the successive elimination of the square roots from the equation $f(\vec{x}) = 0$, we switch from the field element $\hat{f} = a + b\sqrt{d} \in \mathbb{K}_k$ (representing f) to the new field element

$$\hat{f}' = \left(a + b\sqrt{d}\right) \cdot \left(a - b\sqrt{d}\right) = a^2 - b^2 d \in \mathbb{K}_{k-1},$$

which is a representing element of an algebraic function $f'(\vec{x})$ containing less square roots than $f(\vec{x})$. Corresponding to the switching $\hat{f} \mapsto \hat{f}'$, we also replace the original equation $f(\vec{x}) = 0$ by the new equation $f'(\vec{x}) = 0$. Then we keep going by continuing this process until we arrive at a field element

$$\hat{g}(\vec{x}) = \frac{P(\vec{x})}{Q(\vec{x})} \in \mathbb{K}_0 = \mathbb{C}(\mathcal{B}) = \mathbb{C}(\vec{x}),$$

which is a complex rational function $\frac{P(\vec{x})}{Q(\vec{x})}$ of the initial variables $\vec{x}$. Of course, here we assume that the polynomials $P(\vec{x})$ and $Q(\vec{x})$ have no non-trivial common divisor. Then the equation $P(\vec{x}) = 0$ just has the meaning that $f(\vec{x})$ becomes zero on some branch of evaluating the algebraic function f.

By using some simple facts from Galois theory, we can easily detect an intimate relationship between the rational function $\hat{g}(\vec{x}) \in \mathbb{K}_0$ just constructed and the product α of all conjugate elements of $\hat{f} \in \mathbb{K}_n$. (The conjugates of $\hat{f}$ (over the base field $\mathbb{K}_0$) are sitting in the normal hull $\overline{\mathbb{K}_n} : \mathbb{K}_0$ of the field extension $\mathbb{K}_n : \mathbb{K}_0$. The element $\alpha \in \mathbb{K}_0$ is obviously the constant term in the normalized minimal polynomial of $\hat{f}$ over the field $\mathbb{K}_0$. For the elementary concepts and facts of Galois theory, the reader is kindly referred to the book [St(1973)].) Namely, by establishing a natural matching between all conjugates of $a^2 - b^2 d$ and the conjugates of the ordered pair $(a + b\sqrt{d},\, a - b\sqrt{d})$, we can easily see that there exists an integer l such that $\hat{g} = \alpha^{2^l}$. (In the case $l < 0$ this should be understood as $\alpha = (\hat{g})^{2^{-l}}$.) Especially, α and $\hat{g}$ have the same irreducible factors both in the numerator and in the denominator.

Hence the lemma follows. □

Lemma 4.8. *Let S_1 and S_2 be two irreducible polynomials in the complex polynomial ring $\mathbb{C}[x_1, \ldots, x_p]$. Suppose that $S_1 = 0 = S_2$ on the $(p-1)$-dimensional, complex analytic submanifold $J \subset \mathbb{C}^p$. Then $S_2 = \lambda S_1$ with some nonzero $\lambda \in \mathbb{C}$, i. e. S_1 and S_2 are essentially the same.*

Proof. This is an easy consequence of Proposition 4.3 (ii) and Corollary 4.16 of [M (1976)]. Indeed, these results imply that the zero set $Z_1 = \{\vec{x} \in \mathbb{C}^p : \; S_1(\vec{x}) = 0\}$ contains a $(p-1)$-dimensional, dense (in the usual topology), complex analytic and connected submanifold $M_1 = Z_1 \setminus \mathrm{Sing}(Z_1)$, the set of smooth points of Z_1. Since the other polynomial S_2 vanishes on the non-empty, open subset J of M_1, we have that $S_2 = 0$ on M_1. Then the density relation $\overline{M_1} = Z_1$ ensures that $S_2 = 0$ on Z_1. Hilbert's Theorem on Zeroes (Nullstellensatz, cf. (1.5) in §1A. of [M(1976)]) asserts that a power of S_2 is divisible by S_1. Using the irreducibility (prime property) of S_1 and S_2 we get the statement of the lemma. □

Remark 4.9. The ring $\mathbb{K}[x_1, \ldots, x_p]$ of polynomials with p indeterminates over the commutative field $\mathbb{K}$ always has the property of unique factorization into irreducible (prime) factors, see Theorem 13 of Chapter I in [Z-S (1958)].

Corollary 4.10. *Let $S_1, \ldots, S_r$ be polynomials from the complex polynomial ring $\mathbb{C}[x_1, \ldots, x_p]$. Then the common solution set of the equations $S_i = 0$ has dimension at least $p-1$ if and only if these polynomials have a common irreducible divisor.*

Proof. Immediately follows from Lemma 4.8 and Remark 4.9.

Lemma 4.11. *Pertaining to the polynomials $P_i(\vec{x}) \in \mathbb{C}[\vec{x}]$, associated to $R_i(\omega) = f_i(\vec{x}(\omega))$ in Lemma 4.2, we assert the following:*

(i) If the algebraic function $f_i(\vec{x})$ is identically zero on a non-empty, open subset B of $\mathbb{R}^{(2\nu+1)N}$, then f_i is algebraically trivial, i. e. $\hat{f}_i = 0$ or, equivalently, $P_i(\vec{x}) \equiv 0$.

(ii) If $C(N) = 3N!/2^{N-1}$ (defined just as in the forthcoming Lemma 4.21), and the symbolic sequence $\Sigma = (\sigma_1, \ldots, \sigma_n)$ is $C(N)$-rich, then for some i $(1 \le i \le s)$

$P_i(\vec{x}) \not\equiv 0$, and for any common irreducible factor $S(\vec{x})$ of the polynomials $P_j(\vec{x})$ $(1 \le j \le s)$ the smooth part $Z \setminus Sing(Z)$ of the zero set $Z = \{\vec{x} | \, S(\vec{x}) = 0\}$ intersects the Zariski open set D of allowed initial data. (For the definition of D, see the definition of $\tilde{\Omega}$ in the previous section.)

Remarks.

1. In the light of the proof of Lemma 4.8, the assertion in (ii) is equivalent to saying that $Z \cap D \neq \emptyset$ or, equivalently, that $\dim_{\mathbb{C}}(X \setminus D) \le (2\nu+1)N - 2$, where $X \subset \mathbb{C}^{(2\nu+1)N}$ is the common zero set

$$X = \{\vec{x} | \, P_i(\vec{x}) = 0 \text{ for } 1 \le i \le s\}$$

of the polynomials P_i.

2. It is a consequence of (ii) that almost every $(\Sigma; \mathcal{A}; \vec{\tau})$-orbit segment is sufficient, provided that Σ is combinatorially rich as required in (ii). This is a weakened version of Key Lemma 4.1 with a codimension-one exceptional set E.

Proof.

(i): Denote the minimal polynomial of $\hat{f}_i$ over the base field $\mathbb{K}_0 = \mathbb{C}(\vec{x})$ by $M(x) = \sum_{j=0}^{k} c_j x^j$, $c_j \in \mathbb{K}_0$. Suppose, on the contrary, that f_i yet vanishes on an open ball B of $\mathbb{R}^{(2\nu+1)N}$. Since f_i fulfills the equation $\sum_{j=0}^{k} c_j (f_i)^j = 0$, the vanishing of f_i on the ball B implies that the element c_0 of the rational function field $\mathbb{K}_0$ vanishes on B, as well. Therefore, $c_0 = 0$ in $\mathbb{K}_0$, so $M(x)$ is divisible by x. The irreducibility of M yields $M(x) = x$, i. e. $\hat{f}_i = 0$. Hence the first part of the lemma follows. $\square$

(ii): The proof of (ii) will be an induction on $N \ge 2$. For $N = 2$ the statement is obviously true. (Actually, in that case the common zero set X of the polynomials P_i has at least two codimensions, and it is contained in D.) Suppose that $N \ge 3$ and (ii) has been proved for $N-1$. We are going to use, first of all, the purely combinatorial Lemma 4.21, the proof of which is independent of any other lemma in this paper. Namely, let the label of particle N and p, q $(1 \le p < q \le n)$ fulfill the properties (i)–(iv) of Lemma 4.21. The existence of such an N, p, and q is stated by the mentioned lemma.

Suppose, to the contrary of (ii) of the present lemma, that $\dim_{\mathbb{C}}(X \setminus D) = (2\nu+1)N - 1$. Let $J \subset X \setminus D$ be a $[(2\nu+1)N-1]$-dimensional, connected, complex analytic submanifold. We may assume that $\left\| v_{i_1}^0 - v_{j_1}^0 \right\|^2 \equiv 0$ on J; the whole proof is analogous (and, therefore, is left to the reader) in the cases $\left\| v_{i_k}^{k-1} - v_{j_k}^{k-1} \right\|^2 \equiv 0$ or $m_{i_k} + m_{j_k} \equiv 0$, $k = 1, \dots, n$.

The advantage of the identity $\left\| v_{i_1}^0 - v_{j_1}^0 \right\|^2 \equiv 0$ (on J) is that it enables us to eliminate one variable, say $\left(v_{i_1}^0\right)_1$, from $\vec{x}$. Let us write the vector $\vec{x}$ in the form $\vec{x} = \left(\vec{z}; \left(v_{i_1}^0\right)_1\right)$. We can then use the variables in $\vec{z}$ as local (complex) coordinates for J.

The expression of $\left(v_{i_1}^0\right)_1$ via the other variables of $\vec{x}$ only contains the field operations and the square root. Therefore, every polynomial $P_i(\vec{x}) \equiv P_i\left(\vec{z}; \left(v_{i_1}^0\right)_1\right)$ may be re-written on J as an algebraic function $g_i(\vec{z})$ merely containing the field operations and the square root. The identical vanishing (on J) of these polynomials

P_i means that the arising algebraic functions $g_i(\vec{z})$ are algebraically trivial in the sense of (i) of the present lemma.

The algebraic triviality of g_i, however, is preserved by the substitution $m_N = 0$. Suppose for a while that $N \neq i_1$, $N \neq j_1$. Using the substitution $m_N = 0$ and the vanishing of all polynomials P_i on J, we get that for the ball system $\{1, \dots, N\}$ with the infinitely light ball N the equation $\left\| v^0_{i_1} - v^0_{j_1} \right\|^2 = 0$ locally implies non-sufficiency. According to Lemma 4.16 (whose proof is independent of the other lemmas!), this means that the statement (ii) of the present lemma is false for $N-1$ balls. This contradiction finishes the inductive step of the proof in the case $N \notin \{i_1, j_1\}$.

Suppose now that $N = j_1$. Then the vanishing of all polynomials P_i on J, the substitution $m_N = 0$, and the application of Lemma 4.16 in the above manner not only proves that (ii) is false for $N-1$ balls, but it actually implies that *all* $(\Sigma'; \mathcal{A}'; \vec{\tau}')$-orbit segments are non-sufficient.

The proof of Lemma 4.11 is now finished. $\square$

Lemma 4.12. *For any 3-tuple $(\Sigma; \mathcal{A}; \vec{\tau})$ the assertion of the key lemma is true if and only if any common irreducible divisor $R(\vec{x})$ of the polynomials $P_i(\vec{x})$ $(i = 1, \dots, s)$ is purely a polynomial of the masses, i. e. has the form $R(\vec{x}) \equiv R(\vec{m})$.*

Proof. Suppose first that the polynomials $P_i(\vec{x})$ $(i = 1, \dots, s)$ have a common irreducible divisor $R(\vec{x})$ that is not purely a polynomial of $\vec{m}$. Then we show that the statement of the key lemma is false for $(\Sigma; \mathcal{A}; \vec{\tau})$. Indeed, write $R(\vec{x})$ in the form

$$R(\vec{x}) \equiv \sum_{\alpha} T_\alpha(\vec{m}) \cdot (\vec{y})^\alpha \,, \tag{4.13}$$

where

$$\vec{y} = \Big((\tilde{q}^0_i)_j \,, \, (v^0_i)_j \Big|\, i = 1, \dots, N; \; j = 1, \dots, \nu \Big)$$

(so that $\vec{x} = (\vec{y}; \vec{m})$), the summation in (4.13) goes on for finitely many multi-indices $\alpha \in \mathbb{Z}_+^{2\nu N}$, and $T_\alpha(\vec{m})$ is a nonzero complex polynomial of $m_1, \dots, m_N$. By our assumption, the sum in (4.13) contains a term with $|\alpha| > 0$. Therefore, for a generic choice of the N-tuple $\vec{m} = (m_1, \dots, m_N)$ the polynomial $R(\vec{y}; \vec{m})$ is a non-constant polynomial of $\vec{y}$ and, being such, its solution set is $(2\nu N - 1)$-dimensional. Thus, according to lemmas 4.2, 4.8 and 4.11 (ii), the assertion of the key lemma is false.

Suppose now that every common irreducible divisor $R(\vec{x})$ of the polynomials $P_i(\vec{x})$ $(i = 1, \dots, s)$ has the form $R(\vec{x}) \equiv R(\vec{m})$. Write the polynomials $P_i(\vec{x})$ in the form $P_i(\vec{x}) \equiv \overline{R}(\vec{m}) \cdot S_i(\vec{x})$ $(i = 1, \dots, s)$, where $\overline{R}(\vec{m})$ is the greatest common divisor of the polynomials $P_i(\vec{x})$. By Corollary 4.10, the simultaneous solution set X of the equations $S_i(\vec{x}) = 0$, $i = 1, \dots, s$ has dimension at most $(2\nu + 1)N - 2$. A standard Fubini-theorem-type argument yields that for a typical selection of masses $\vec{m}$ the section $X_{\vec{m}}$ of X (corresponding to $\vec{m}$) has at least two codimensions in $\mathbb{C}^{2\nu N}$. On the other hand, for a generic $\vec{m}$ one also has $\overline{R}(\vec{m}) \neq 0$. Therefore, the $\vec{m}$-section $Y_{\vec{m}}$ of the common solution set Y of $\overline{R}(\vec{m}) \cdot S_i(\vec{x}) = 0$ has codimension at least two in $\mathbb{C}^{2\nu N}$, and this is just the assertion of the key lemma. Hence Lemma 4.12 follows. $\square$

Lemma 4.14. *The polynomials $P_i(\vec{x})$ $(i = 1, \dots, s)$ are homogeneous in the masses $\vec{m}$.*

Proof. The equations (3.4), (3.7)–(3.9) (which are equivalent to (3.4)–(3.8)) are all homogeneous with degree zero in $\vec{m}$. In other words, the whole complex dynamics is invariant under the rescalings of masses $\vec{m} \mapsto \lambda \cdot \vec{m}$ $(0 \neq \lambda \in \mathbb{C})$ or, equivalently, the dynamics only depends on the ratios m_i/m_N, $i = 1, \dots, N-1$. Therefore, the algebraic functions $f_i(\vec{x})$ are also homogeneous with degree zero in $\vec{m}$, so the polynomials $P_i(\vec{x})$ must also be homogeneous in $\vec{m}$. $\square$

Corollary 4.15. *Every factor of the polynomial $P_i(\vec{x})$ is homogeneous in $\vec{m}$.*

Proof. Any factor a homogeneous polynomial is easily seen to be also homogeneous. $\square$

The next lemma describes the $(\Sigma; \mathcal{A}; \vec{\tau}; \vec{m})$-orbit segments in the case when the N-th ball is "infinitely light" compared to the others, i. e. when $m_N = 0$ but $m_1 \cdot \ldots \cdot m_{N-1} \neq 0$.

Lemma 4.16. *Suppose that $m_N = 0$ but $m_1 \cdot \ldots \cdot m_{N-1} \neq 0$. Assume further that there exist indices $1 \leq p < q \leq n$ such that $N \in \sigma_p$, $N \in \sigma_q$, $N \notin \sigma_j$ for $j = p+1, \dots, q-1$, and, finally, if $\sigma_p = \sigma_q$ then we require additionally that there exists an index j, $p < j < q$, such that $\sigma_p \cap \sigma_j \neq \emptyset$. (In the vocabulary of [K-S-Sz(1992)] this means that σ_p and σ_q do not belong to the same "island".) Then an orbit segment $\omega \in \tilde{\Omega}(\Sigma; \mathcal{A}; \vec{\tau}; \vec{m})$ is sufficient (i. e. $dim_{\mathbb{C}} \mathcal{N}(\omega) = \nu + 1$) if*

(1) the $\{1, 2, \dots, N-1\}$-part

$$\left\{\tilde{q}_i^k(\omega),\, v_i^k(\omega) \,|\, i = 1, \dots, N-1;\ k = 0, \dots, n\right\}$$

of the orbit segment is sufficient as an orbit segment of the particles $1, \dots, N-1$; and

(2) the relative velocities $v_N^p(\omega) - v_{i_p}^p(\omega)$ and $v_N^{q-1}(\omega) - v_{i_q}^{q-1}(\omega)$ are not parallel.

(Here, as usual, i_p (i_q) is the index of the ball colliding with the N-th particle at σ_p (σ_q).)

Remark. It is worth noting that, as it follows easily from the conditions, $v_N^p(\omega) = v_N^{q-1}(\omega)$.

Proof. First of all, we observe that the motion of the N-th ball with zero mass has absolutely no effect on the evolution of the $\{1, \dots, N-1\}$-part of the trajectory segment, and this statement is also valid for the time evolution of the $\{1, \dots, N-1\}$-part $\left\{\delta\tilde{q}_i^k \,|\, i = 1, \dots, N-1\right\}$ of a neutral vector $\left(\delta\tilde{q}_1^0, \dots, \delta\tilde{q}_N^0\right) \in \mathcal{N}(\omega)$, see also the equations (3.4), (3.7)–(3.8) and (3.10). We note here that the reflection laws for the neutral vectors are the same as for velocities.

Consider now an arbitrary neutral vector $\delta Q = \left(\delta\tilde{q}_1^0, \dots, \delta\tilde{q}_N^0\right) \in \mathcal{N}(\omega)$. According to (1) and the above principle, we can modify the neutral vector δQ by a scalar multiple of the flow direction and by a uniform spatial displacement in such a way that $\delta\tilde{q}_i^0 = 0$ for $i = 1, \dots, N-1$, and then $\delta\tilde{q}_i^k = 0$ remains true for the whole orbit segment, $i = 1, \dots, N-1$, $k = 0, \dots, n$. The neutrality of δQ with respect to the entire orbit segment means, however, that

$$\begin{aligned} \delta\tilde{q}_N^p &= \alpha_p \cdot \left(v_N^p(\omega) - v_{i_p}^p(\omega)\right) \\ \delta\tilde{q}_N^{q-1} &= \alpha_q \cdot \left(v_N^{q-1}(\omega) - v_{i_q}^{q-1}(\omega)\right). \end{aligned} \tag{4.17}$$

According to our hypothesis (2), the consequence of (4.17) and of the obvious equation $\delta\tilde{q}_N^p = \delta\tilde{q}_N^{q-1}$ is that $\alpha_p = \alpha_q = 0$, and thus $\delta Q = 0$. Hence the lemma follows. □

Lemma 4.18. *Keep the notations and conditions of the previous lemma except hypotheses (1) and (2). Then hypothesis (2) is violated only on a subset A of $\tilde{\Omega}(\Sigma; \mathcal{A}; \vec{\tau}; \vec{m})$ with codimension at least two.*

Proof. We are applying the complex version of Lemma 3.25 of [Sim(1992)-II] to the subsystem $\{1, \dots, N-1\}$. It follows from that result that the codimension of A in $\tilde{\Omega}$ is at least $\nu - 1$, and this number is at least two in our setup of $\nu \geq 3$. Hence the lemma follows. □

Remark. Note that the proof of the above lemma is the single point in the whole paper where we take advantage of the condition $\nu \geq 3$.

Lemma 4.19. *Assume that the combinatorial scheme $(\Sigma; \mathcal{A}; \vec{\tau})$ is given for the N (≥ 3) ball system $\{1, \dots, N\}$ and, furthermore, there are two integers $1 \leq p < q \leq n$ such that $N \in \sigma_p$, $N \in \sigma_q$, $N \notin \sigma_j$ for $p < j < q$, and if $\sigma_p = \sigma_q$ then we require additionally that there exists an index j, $p < j < q$, such that $\sigma_p \cap \sigma_j \neq \emptyset$. Denote by $(\Sigma'; \mathcal{A}'; \vec{\tau}')$ the sub-scheme of the above one that can be obtained by discarding all σ_j, a_j, and τ_j corresponding to the collisions σ_j involving the ball N. (This sub-scheme describes the time evolution of the $\{1, \dots, N-1\}$-subsystem when having $m_N = 0$.)*

Denote by $P_1(\vec{x}), \dots, P_s(\vec{x})$ the complex polynomials, the simultaneous vanishing of which means the non-sufficiency of the $(\Sigma; \mathcal{A}; \vec{\tau})$-orbit segments in the sense of Lemma 4.2. We claim that if these polynomials have a common irreducible divisor, then the similar polynomials associated with the sub-scheme $(\Sigma'; \mathcal{A}'; \vec{\tau}')$ also have a common irreducible factor.

Proof. Suppose that the polynomials $P_1(\vec{x}), \dots, P_s(\vec{x})$ are divisible by the irreducible polynomial $R(\vec{x})$: $P_i(\vec{x}) \equiv R(\vec{x}) \cdot T_i(\vec{x})$. Now substitute in these polynomial identities $m_N = 0$ and get

$$P_i^*(\vec{y}; \vec{m}') \equiv R^*(\vec{y}; \vec{m}') \cdot T_i^*(\vec{y}; \vec{m}'), \tag{4.20}$$

$i = 1, \dots, s$, where $\vec{m}' = (m_1, \dots, m_{N-1})$ and the polynomials with superscript $*$ were obtained from the "starless" ones by the mentioned substitution. The polynomial $R^*(\vec{y}; \vec{m}')$ is a divisor of $P_i^*(\vec{y}; \vec{m}')$, $i = 1, \dots, s$. If $R^*(\vec{y}; \vec{m}') \equiv 0$, then $P_i^*(\vec{y}; \vec{m}') \equiv 0$ for $i = 1, \dots, s$, and we are done according to Lemma 4.16. (Then, by Lemma 4.11 (i), the polynomials $P_i'(\vec{y}'; \vec{m}')$ describing the non-sufficiency of the orbit segment of the $(\Sigma'; \mathcal{A}'; \vec{\tau}')$ subsystem are all identically zero.) Thus, we can assume that $R^*(\vec{y}; \vec{m}') \not\equiv 0$.

The polynomial $R^*(\vec{y};\vec{m}')$ cannot be constant. Indeed, according to Corollary 4.15, the polynomial $R(\vec{y};\vec{m})$ is homogeneous in the masses. If the degree of homogeneity of R is zero, then $R^* \equiv R$ and R^* is not constant (being irreducible). If, on the other hand, the mentioned degree is positive, then every monomial of R^* is divisible by some m_i $(i < N)$, so R^* is again non-constant.

Consider all $(\Sigma;\mathcal{A};\vec{\tau})$ orbit segments with the side condition $m_N = 0$. The initial data of such orbit segments form an open and dense subset of $\mathbb{C}^{(2\nu+1)N-1}$. Using the assertion (ii) of Lemma 4.11 (More precisely: using the analogue of (ii) of Lemma 4.11 in the case $m_N = 0$.) we see that the zero set of the non-constant polynomial $R^*(\vec{y};\vec{m}')$ provides a $((2\nu+1)N-2)$-dimensional family of orbit segments that are not sufficient. According to Lemma 4.16, this means that there is a codimension-one family of $(\Sigma';\mathcal{A}';\vec{\tau}')$ orbit segments that are not sufficient. Hence, by Corollary 4.10, the polynomials $P_i'(\vec{y}';\vec{m}')$ (describing the non-sufficiency of the $(\Sigma';\mathcal{A}';\vec{\tau}')$-dynamics) must have a common ireducible divisor. The lemma is now proven. □

The next, purely combinatorial lemma is the last ingredient of the inductive proof of the key lemma.

Lemma 4.21. *Define the sequence of positive numbers $C(N)$ recursively by taking $C(2) = 1$ and $C(N) = \frac{N}{2} \cdot max\{C(N-1);3\}$ for $N \geq 3$. Let $N \geq 3$, and suppose that the symbolic collision sequence $\Sigma = (\sigma_1,\ldots,\sigma_n)$ for N particles is $C(N)$-rich. Then we can find a particle, say the one with label N, and two indices $1 \leq p < q \leq n$ such that*

(i) $N \in \sigma_p \cap \sigma_q$,

(ii) $N \notin \bigcup_{j=p+1}^{q-1} \sigma_j$,

(iii) $\sigma_p = \sigma_q \Longrightarrow \exists j\ p < j < q\ \&\ \sigma_p \cap \sigma_j \neq \emptyset$,

and

(iv) Σ' is $C(N-1)$-rich on the vertex set $\{1,\ldots,N-1\}$.

Here, just as in Lemma 4.19, we denote by Σ' the symbolic sequence that can be obtained from Σ by discarding all edges containing N.

Proof. The hypothesis on Σ implies that there exist subsequences $\Sigma_1,\ldots,\Sigma_r$ of Σ with the following properties:

(1) For $1 \leq i < j \leq r$ every collision of Σ_i precedes every collision of Σ_j,

(2) the graph of Σ_i $(1 \leq i \leq r)$ is a tree (a connected graph without loop) on the vertex set $\{1,\ldots,N\}$,

and

(3) $r \geq C(N)$.

Since every tree contains at least two vertices with degree one and $C(N) = \frac{N}{2} \cdot \max\{C(N-1);3\} = \frac{N}{2} \cdot b$ (b is a shorthand for $\max\{C(N-1);3\}$) we have that there is a vertex, say the one labelled by N, such that N is a degree-one vertex of $\Sigma_{i(1)},\ldots,\Sigma_{i(t)}$, where $1 \leq i(1) < \cdots < i(t) \leq r$ and $t \geq b$.

Now $t \geq C(N-1)$ and, therefore, (iv) obviously holds.

Let $\sigma_{p'}$ the edge of $\Sigma_{i(1)}$ that contains N and, similarly, let $\sigma_{q'}$ be the edge of $\Sigma_{i(t)}$ containing the vertex N. Then the fact $t \geq 3$ ensures that the following properties hold:

(i)' $N \in \sigma_{p'} \cap \sigma_{q'}$,

(iii)' $\sigma_{p'} = \sigma_{q'} \Longrightarrow \exists j\ p' < j < q'\ \&\ \sigma_{p'} \cap \sigma_j \neq \emptyset,\ \sigma_j \neq \sigma_{p'}$.

Let σ_p, σ_q $(1 \le p < q \le n)$ be a pair of edges $\sigma_{p'}$, $\sigma_{q'}$ $(1 \le p' < q' \le n)$ fulfilling (i)' and (iii)' and having the minimum possible value of $q' - p'$. Elementary inspection shows that then (ii) must also hold for σ_p, σ_q. Lemma 4.21 is now proved. □

The Inductive Proof of Key Lemma 4.1

We prove a bit stronger result (cf. also Lemma 4.12):

If the numbers $C(N)$ $(N \ge 2)$ are defined as in Lemma 4.21 and all hypotheses of the key lemma are fulfilled by Σ, then the polynomials $P_i(\vec{x})$ (associated with $(\Sigma; \mathcal{A}; \vec{\tau})$ by Lemma 4.2) have no common irreducible factor.

Indeed, for $N = 2$ there are no non-sufficient trajectories with initial data in D and, therefore, the above statement is obviously true.

The inductive step $(N-1) \to N$ can be obtained by simply putting together Lemmas 4.21 and 4.19.

The proof of the key lemma is now complete in the complex case. □

Finally, the following, very simple lemma from complex analysis (a local statement), together with Lemma 4.8, takes care of the real version of Key Lemma 4.1.

Lemma 4.22. *Let $p \in \mathbb{N}$ and $f : \mathbb{C}^p \to \mathbb{C}$ be a holomorphic function defined in a neighborhood U of the origin of $\mathbb{C}^p$. Assume that*

(1) $f(0) = 0$,

(2) $Df(0) \neq 0$, and

(3) all Taylor coefficients of f at 0 are real.

Consider the complex analytic submanifold

$$J = \{x \in U \,|\, f(x) = 0\}$$

of U having a $(p-1)$-dimensional, completely real intersection with $\mathbb{R}^p$. (We can assume – by choosing the shape and smallness of U appropriately – that both J and $J \cap \mathbb{R}^p$ are connected.)

We assert that if a holomorphic function $g : J \to \mathbb{C}$ vanishes on $J \cap \mathbb{R}^p$, then $g = 0$ on the entire J.

Proof. It follows from (1)-(3) that the derivative of the projection of J onto the subspace spanned by some $p-1$ coordinate axes, say by $z_1, \dots, z_{p-1}$, has maximum rank $p-1$ at the origin. Then $z_1, \dots, z_{p-1}$ may be chosen as holomorphic local coordinates in J near zero. The condition $\left(g\big|J \cap \mathbb{R}^p\right) \equiv 0$ implies that all Taylor coefficients of g in this local coordinate system are zero and, therefore, $g = 0$ on J. □

The proof of the real version of Key Lemma 4.1 is now complete. □

5. Proof of the Main Theorem

Quite until the last Remark, throughout the whole section the dimension $\nu \geq 3$ of the system will be fixed.

Let us fix a system satisfying the conditions of the Main Theorem where the number of particles is $N \geq 2$. The proof of the K-property of the standard billiard flow with mass vector $\vec{m}$ will go on by induction on the number N of particles. Assume thus that the standard billiard ball system with mass vector $\vec{m}$ has the K-property for any number of particles $n < N$ whenever the interior of the configuration space is connected — apart from a countable union of proper analytic submanifolds of mass vectors $(m_1, \dots, m_n)$. For $n = 2$ this is the case since the standard billiard flow $(\mathbf{M}, S^{\mathbb{R}}, \mu)_{\vec{m}})$ is a dispersing billiard, and [S-Ch (1987)] is applicable.

It is important to note here that, in fact, the Bernoulli property of the standard billiard flow $(\mathbf{M}, \{S^t\}, \mu)$ follows from its weak mixing, as it has been recently shown by Chernov–Haskell and Ornstein–Weiss, cf. [C-H(1996)] and [O-W(1996)]. For a brief summary of how the K-property follows, see – for instance – Section 7 of [S (1992)-I] , or [K-S(1986)].

Richness will be understood in the sense of Defintion 2.5 by choosing $C := C(N)$ as prescribed by Key Lemma 4.1 (see the remark after the lemma). We stress again that the proof of the key lemma is independent of the inductive assumption.

The whole strategy outlined in Section 1 is based on the Theorem of Local Ergodicity for semi–dispersing billiards. In fact, the proof of global ergodicity uses its form given in [S-Ch (1987)] , whereas the inductive proof of part of the big proof requires its more general, transversal form as phrased in [K-S-Sz (1990)] . For an easier identification of its conditions we refer to the work [K-S-Sz (1990)] .

According to the results of [V (1979)], [G (1981)] and [B-F-K(1995)], in a semi-dispersing billiard system there are no orbits with a finite accumulation point of collision moments. Consequently, Condition 2.1 of [K-S-Sz (1990)] is fulfilled automatically. Condition 3.1, the so-called Chernov-Sinai Ansatz, will be verified at the end of the section. Condition 3.2 on the regularity of the set of degenerate tangencies is automatic for cylindric billiards. The argument given in [K-S-Sz (1991)] for Condition 3.3 on the regularity of the set of double singularities for the case $N = 3$ also works in our general case. What is more, Lemma 4.15 of [K-S-Sz (1991)] asserts that the set of orbits containing two or more singular collisions is a countable union of codimension-2 submanifolds of $\mathbf{M}$. Thus, it is sufficient to concentrate our proof on the set $\mathbf{M}^*$ of orbits containing at most one singular collision, and to show that, outside a slim set, any $x \in \mathbf{M}^*$ has an open neighbourhood belonging to one ergodic component. This is what we will do.

Return first to the fundamental theorem (Theorem on Local Ergodicity, Theorem 5 of [S-Ch (1987)] or Corollary 3.12 of [K-S-Sz (1990)]) which is applicable to our system of N particles, saying that *a suitably small, open neighborhood of any sufficient phase point with at most one singularity belongs to one ergodic component.*

Since the complement of a slim set in a connected manifold is always connected (see Property 2.10 from [K-S-Sz (1991)]), our goal is to show that, apart from a slim set of points, every phase point with at most one singularity is sufficient. To do so, let us follow our three step strategy outlined in the introduction. Its non-inductive part (1) is settled by the real version of Key Lemma 4.1 for typical mass vectors, at least. (The phrase *typical* means that"apart from a countable union of proper analytic submanifolds of mass vectors $\vec{m} \in \mathbb{R}_+^N$".) For step (2) we use the following generalization of Theorem 5.1 of [S (1992)-I] to particles with different masses.

Propostion 5.1. *Consider a system of N (≥ 3) particles on the unit torus $\mathbb{T}^\nu$ $(\nu \geq 2)$ satisfying $r \in R_0$. Let $P = \{P_1, P_2\}$ be a given, two–class partition of the N particles, where, for simplicity, $P_1 = \{1, \dots, n\}$ and $P_2 = \{n+1, \dots, N\}$ $(n < N-1)$. Suppose further that the standard billiard flows with mass vectors $\vec{m}_{P_1} = (m_1, \dots, m_n)$ and $\vec{m}_{P_2} = (m_{n+1}, \dots, m_N)$ are K-systems (if $n = 1$, then the K-property is only assumed about the second subsystem). Then the set*

$$F_+ = \left\{x \in \mathbf{M} : S^{[0,\infty)}x \text{ is partitioned by } P\right\}$$

is a closed, zero set with codimension at least two (i.e. a closed slim set).

This statement would, in fact, allow to fix $C(N)$ (the number of consecutive, connected collision graphs required for sufficiency) arbitrarily large.

Proof. The reader familiar with the proof of Theorem 5.1 of [S (1992)-I] will immediately see that the same arguments work here, too, irrespective of the fact that the notion of center of mass has lost its sense now. The most direct way to convince ourselves about this is the following one: when constructing the configuration space of our system in Section 1, we were using the equivalence relation Ψ. Thus the elements of the configuration space $\mathbf{Q}$ can be represented in the following way: we fix, for instance, $q_1 = 0$, and then a configuration in $\mathbf{Q}$ can be given by all the relative positions $q_j - q_1 :\ 2 \leq j \leq N$. It is easy to see that we are allowed to use this representation for the configuration spaces $\mathbf{Q}$ and $\mathbf{Q}_1$, but then in $\mathbf{Q}_2$ we are only permitted to use a similar representation by fixing, for instance, $q_{n+1} = a$. The position a then determines the relative initial configurations of the subsystems corresponding to the partition P in the same way as the center of mass $C_1(x)$ did in the aforementioned proof of [S (1992)-I] , and consequently, the arguments can, indeed, be repeated without any change. $\square$

Consider now a rich point $x \in \mathbf{M}^0$. (By Proposition 5.1, we can assume the richness of x.) If $x \in \mathbf{M}^0$ is not sufficient, then by Key Lemma 4.1, there is an open neighbourhood U of x and a CW-complex (more accurately, a closed algebraic subset) $\mathcal{N} \subset U$, $x \in \mathcal{N}$, such that $\operatorname{codim}_U \mathcal{N} \geq 2$, and all non-sufficient points of U are contained in $\mathcal{N}$. Consequently, by the Theorem on Local Ergodicity and by the connectedness of the set $U \setminus \mathcal{N}$, the neighborhood U belongs to one ergodic component (modulo zero sets, of course).

Let us finally handle the phase points $x \in \mathbf{M}^1$. We can, and do assume that x belongs, for instance, to the set $\mathcal{SR}^+$. (The set $\mathcal{SR}^+$ is the collection of all phase points $x \in \partial\mathbf{M}$ for which the reflection, occurring at x, is singular (tangential or multiple) and, in the case of a multiple collision, x is supplied with the *outgoing* velocity V^+.)

Here the following generalization of Theorem 6.1 of [S (1992)-I] to systems of particles with different masses will be of help.

Proposition 5.2. *Consider a system of N (≥ 3) particles on the unit torus $\mathbb{T}^\nu$ ($\nu \geq 2$) satisfying $r \in R_0$. Let $P = \{P_1, P_2\}$ be a given, two–class partition of the N particles, where, for simplicity, $P_1 = \{1, \dots, n\}$ and $P_2 = \{n+1, \dots, N\}$ ($n < N-1$). Suppose further that the standard billiard flows with mass vectors $\vec{m}_{P_1} = (m_1, \dots, m_n)$ and $\vec{m}_{P_2} = (m_{n+1}, \dots, m_N)$ are K-systems (if $n = 1$, then, of course, the K-property is only assumed about the second subsystem). Suppose, moreover, that Key Lemma 4.1 has been proved for billiard flows satisfying the conditions of the Main Theorem. Then, for any positive number t_0 and any cell $C \subset \mathcal{SR}^+$ of $\mathcal{SR}^+$ with the maximum dimension $(2d-3)$ the subset*

$$C(t_0, P) = \left\{ x \in C : \; S^{[t_0, \infty)} x \text{ is partitioned by } P \right\}$$

has measure zero with respect to the Riemannian volume μ_C of the manifold C.

Proof. By using the remarks made in the proof of Proposition 5.1, the arguments of the proof of Theorem 6.1 of [S (1992)-I] can be adapted to provide our more general statement, too. □

Note that - since the set of doubly singular phase points is a countable union of codimension-two submanifolds, as said before - the condition of regularity of the orbit $S^{[0,\infty)}x$ can be omitted from the definition of $C(t_0, P)$. Since the set $C(t_0, P)$ in this theorem is closed and, therefore, it is a *slim* subset of $\mathbf{M}$, we may assume that x does not belong to any of these sets, i. e. the (non-singular) positive semi-trajectory $S^{(0,\infty)}x$ contains infinitely many consecutive, connected collision graphs. By Key Lemma 4.1, this positive orbit is sufficient, if only x does not belong to some codimension-one submanifolds of $\mathcal{SR}^+$, but such sets are slim (negligible) in the whole phase space $\mathbf{M}$. In this way we have proved two things:

(a) the Chernov-Sinai Ansatz,

and the fact that

(b) apart from a slim subset of the phase space, the trajectory of every phase point $x \in \mathbf{M}^1$ is sufficient.

The inductive proof of the Main Theorem is now complete. □

Remark. Consider finally the case $\nu = 2$ (cf. Note 1 of Section 1). The careful reader has certainly observed that throughout the whole paper the assumption $\nu \geq 3$ was only used in Lemma 4.18 when applying Lemma 3.25 of [S (1992)-II] . In the case $\nu = 2$ this particular lemma provides that the codimension of the subset $A \subset \tilde{\Omega}$ occuring in our Lemma 4.18 is at least 1. Consequently, by combining the ideas of the proof for the case $\nu \geq 3$ with those used in proving Theorems C and D in [S-Sz(1995-II)] we obtain that the ergodic components of the system are, indeed, open as claimed in the aforementioned Note 1.

6. Concluding remarks.

1. The system of $N = 2$ hard discs (i. e. $\nu = 2$) with arbitrary masses m_1, m_2 was studied by Simányi and Wojtkowski [S-W(1989)]. They were considering the

same dynamics as here but in the larger phase space $\tilde{\mathbf{Q}} \times S^1$. By using the fact that it is an isometric $\mathbb{T}^2$-extension of the standard billiard ball flow $(\mathbf{M}, S^{\mathbb{R}}, \mu)_{(m_1, m_2)}$, which is a planar dispersing billiard and was known to possess the K-property since [S(1970)], they showed that this extension is also a K-flow as long as the ratio m_1/m_2 is irrational. We expect that by combining the methods of the present paper with their approach, our results can also be extended to obtain the K-property of the flow on $\tilde{\mathbf{Q}} \times \mathcal{E}$ from that on $\mathbf{Q} \times \mathcal{E}$.

2. The inductive steps of the present paper used the fact that, when $m_N \to 0$, then the collision laws extend analytically to the case $m_N = 0$, and, moreover, the motion of the subsystem of balls $\{1, \dots, N-1\}$ gets independent of that of ball N. Thus for the dynamics of this subsystem we were able to use the inductive assumption.

We note that the case is similar when $m_N \to \infty$: the N-th ball of infinitely large mass becomes a fixed ball-shaped scatterer, and again an inductive assumption about $N-1$ moving particles among a finite number of fixed ball-shaped scatterers can be used. In this way, part (1) of the strategy outlined in the Introduction can be settled easily. We expect that the inductive statements used to prove parts (2) and (3) can be generalized to this model, too, and consequently, the K-property of these models can also be established for typical mass vectors, at least.

3. In [Sz(1994)], the second author obtained necessary and sufficient conditions for the K-property of so-called "orthogonal cylindric billiards". The subclass of cylindric billiards, introduced in [Sz(1993)], within the family of semi-dispersing billiards is of particular interest since it is this subclass, for which the formulation of general — and at the same time in some sense constructively verifiable — necessary and sufficient conditions of the K-property seems possible. (We have seen in the proof of Lemma 2.1 that the system of N balls with an arbitrary mass vector is also isomorphic to a cylindric billiard.) Indeed, such a condition is found and shown to be necessary for the systems of N hard balls of identical masses in a forthcoming paper of the present authors [S-Sz(1996)], which is based upon the ideas from [S-Sz(1995-I)]. One apparent difficulty in establishing the sufficiency of this condition is the fact that so far no characterization of the neutral subspace of a trajectory segment of a general cylindric billiard is known (in the case of hard ball systems such a characterization was obtained by using the CPF, whose derivation, on the other hand, used the conservation of momentum, a property present in hard ball systems but absent in general cylindric billiards).

Acknowledgement. Special thanks are due to János Kollár who called our attention to the advantage of the complexification of the dynamics which turned out to be very important in sections 3-4. Also, he made useful remarks concerning some elementary facts from algebraic geometry. The authors express their sincere gratitude to N. Chernov, A. Katok, C. Liverani, Ya. Pesin and L. Vaserstein for their helpful remarks on earlier versions of the paper.

References

[B-F-K(1995)] D. Burago, S. Ferleger, A. Kononenko, *Uniform estimates on the number of collisions in semi-dispersing billiards.*, Manuscript..

[B-S(1973)] L. A. Bunimovich, Ya. G. Sinai, *The fundamental theorem of the theory of scattering billiards*, Math. USSR-Sb. **19**, 407-423 (1973).

[C-F-S(1982)] I. P. Cornfeld, S. V. Fomin, Ya. G. Sinai, *Ergodic Theory*, Berlin, Heidelberg, New York: Springer 1982.

[C-H(1996)] N. I. Chernov, C. Haskell, *Nonuniformly hyperbolic K-systems are Bernoulli*, Ergodic Theory and Dynamical Systems **16**, 19-44 (1996).

[G(1981)] G. Galperin, *On systems of locally interacting and repelling particles moving in space*, Trudy MMO **43**, 142-196 (1981).

[H-W(1941)] W. Hurewicz, H. Wallman, *Dimension Theory*, Princeton, 1941.

[K(1979)] N. S. Krylov, *The Processes of Relaxation of Statistical Systems and the Criterion of Mechanical Instability*, Thesis, Moscow, (1942); *Republished in English by Princeton University Press*, Princeton N. J. (1979).

[K-S(1986)] A. Katok, J.-M. Strelcyn, *Invariant Manifolds, Entropy and Billiards; Smooth Maps with Singularities*, Lecture Notes in Mathematics **1222**, Springer, 1986.

[K-S-Sz(1989)] A. Krámli, N. Simányi, D. Szász, *Ergodic Properties of Semi–Dispersing Billiards I. Two Cylindric Scatterers in the 3–D Torus*, Nonlinearity **2**, 311–326 (1989).

[K-S-Sz(1990)] A. Krámli, N. Simányi, D. Szász, *A "Transversal" Fundamental Theorem for Semi-Dispersing Billiards*, Commun. Math. Phys. **129**, 535–560 (1990).

[K-S-Sz(1991)] A. Krámli, N. Simányi, D. Szász, *The K–Property of Three Billiard Balls*, Annals of Mathematics **133**, 37–72 (1991).

[K-S-Sz(1992)] A. Krámli, N. Simányi, D. Szász, *The K–Property of Four Billiard Balls*, Commun. Math. Phys. **144**, 107-148 (1992).

[M(1976)] D. Mumford, *Algebraic Geometry I. Complex Projective Varieties*, Springer-Verlag (1976).

[O-W(1996)] D. Ornstein, B. Weiss, *On the Bernoulli Nature of Systems with Some Hyperbolic Structure*, Ergodic Theory and Dynamical Systems (accepted for publication), pp. 23 (1996).

[Sin(1963)] Ya. G. Sinai, *On the Foundation of the Ergodic Hypothesis for a Dynamical System of Statistical Mechanics*, Soviet Math. Dokl. **4**, 1818-1822 (1963).

[Sin(1970)] Ya. G. Sinai, *Dynamical Systems with Elastic Reflections*, Russian Math. Surveys **25:2**, 137-189 (1970).

[St(1973)] I. Stewart, *Galois Theory*, Chapman and Hall, London (1973).

[Sin(1979)] Ya. G. Sinai, *Development of Krylov's Ideas*, Afterword to N. S. Krylov's "Works on the foundations of statistical physics", see reference [K(1979)], Princeton University Press (1979).

[Sim(1992)] N. Simányi, *The K-property of N billiard balls I*, Invent. Math. **108**, 521-548 (1992); *II.*, ibidem **110**, 151-172 (1992).

[S-Ch(1987)] Ya. G. Sinai, N.I. Chernov, *Ergodic properties of certain systems of 2–D discs and 3–D balls*, Russian Math. Surveys **(3) 42**, 181-207 (1987).

[S-Sz(1994)] N. Simányi, D. Szász, *The K-property of 4-D Billiards with Non-Orthogonal Cylindric Scatterers*, J. Stat. Phys. **76, Nos. 1/2**, 587-604 (1994).

[S-Sz(1995)-I] N. Simányi, D. Szász, *The Boltzmann-Sinai ergodic hypothesis for hard ball systems*, Manuscript.

[S-Sz(1995)-II] N. Simányi, D. Szász, *The K-property of Hamiltonian systems with restricted hard ball interactions*, Mathematical Research Letters **2, No. 6**, 751-770 (1995).

[S-Sz(1996)] N. Simányi, D. Szász, *Cylindric Billiards and Transitive Lie Group Actions*, in preparation.

[S-W(1989)] N. Simányi, M. Wojtkowski, *Two-particle billiard system with arbitrary mass ratio*, Ergod. Th. & Dynam. Sys. **9**, 165-171 (1989).

[Sz(1993)] D. Szász, *Ergodicity of classical billiard balls*, Physica A **194**, 86-92 (1993).

[Sz(1994)] D. Szász, *The K-property of 'Orthogonal' Cylindric Billiards*, Commun. Math. Phys. **160**, 581-597 (1994).

[Sz(1996)] D. Szász, *Boltzmann's Ergodic Hypothesis, a Conjecture for Centuries?*, Studia Sci. Math. Hung **31**, 299-322 (1996).

[VDW(1955)] B. L. Van Der Waerden, *Algebra I*, Springer-Verlag, 1955.

[V(1979)] L. N. Vaserstein, *On Systems of Particles with Finite Range and/or Repulsive Interactions*, Commun. Math. Phys. **69**, 31-56 (1979).

[Z-S(1958)] O. Zariski, P. Samuel, *Commutative Algebra I.*, D. Van Nostrand Company (1958).