

MYSTERY OF POINT CHARGES

A. GABRIELOV, D. NOVIKOV AND B. SHAPIRO

To Vladimir Igorevich Arnold who taught us to study classics

ABSTRACT

We discuss the problem of finding an upper bound for the number of equilibrium points of a potential of several fixed point charges in \mathbb{R}^n . This question goes back to J. C. Maxwell and M. Morse. Using fewnomial theory we show that for a given number of charges there exists an upper bound independent of the dimension, and show it to be at most 12 for three charges. We conjecture an exact upper bound for a given configuration of non-negative charges in terms of its Voronoi diagram, and prove it asymptotically.

1. Introduction

Consider a configuration of $l = \mu + \nu$ fixed point charges in \mathbb{R}^n , with $n \geq 3$, consisting of μ positive charges with the values $\zeta_1, \dots, \zeta_\mu$, and ν negative charges with the values $\zeta_{\mu+1}, \dots, \zeta_l$. They create an electrostatic field whose potential equals

$$V(\bar{x}) = \left(\frac{\zeta_1}{r_1^{n-2}} + \dots + \frac{\zeta_\mu}{r_\mu^{n-2}} \right) + \left(\frac{\zeta_{\mu+1}}{r_{\mu+1}^{n-2}} + \dots + \frac{\zeta_l}{r_l^{n-2}} \right), \quad (1.1)$$

where r_i is the distance between the i th charge and the point $\bar{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ which we assume to be different from the locations of the charges. Below we consider the problem of *finding effective upper bounds* on the number of critical points of $V(\bar{x})$, that is, the number of points of equilibrium of the electrostatic force. In what follows we mostly assume that considered configurations of charges have only non-degenerate critical points. This guarantees that the number of critical points is finite. Such configurations of charges and potentials will be called *non-degenerate*. Surprisingly little is known about this whole topic and the references are very scarce.

The classical case of $\mathbb{R}^2 \simeq \mathbb{C}$ when the potential of the unit charge placed at the origin equals $\ln|z|$ was a subject of investigation by C. F. Gauss; see [4, 11]. In particular, he noticed that the electrostatic force, that is, the gradient of $V(\bar{x}) = \sum \zeta_i \log|z - z_i|$, created by a configuration of charges with values ζ_1, \dots, ζ_l located at the points z_1, \dots, z_l respectively equals

$$F(z) = \sum_{i=1}^l \frac{\zeta_i}{z - z_i}$$

which implies that the equilibrium points of this force coincide with the zeros of the rational function in the right-hand side of the latter expression. Therefore, their total number is at most l .

In the case of \mathbb{R}^3 one of the very few known results obtained by direct application of Morse theory to $V(\bar{x})$ is as follows; see [14, Theorem 32.1; 8, Theorem 6].

Received 10 March 2005; revised 14 June 2006; published online 1 June 2007.

2000 *Mathematics Subject Classification* 31B05 (primary), 58E05 (secondary).

The research of the first author was supported by NSF grants DMS-0200861 and DMS-0245628. That of the second author was supported by NSF grant DMS-0200861.

THEOREM 1.1 (Morse and Kiang). *Assume that the total charge $\sum_{i=1}^l \zeta_i$ in (1.1) is negative (respectively positive). Let m_1 be the number of the critical points of index 1 of V , and m_2 be the number of the critical points of index 2 of V . Then $m_2 \geq \mu$ (respectively $m_2 \geq \mu - 1$) and $m_1 \geq \nu - 1$ (respectively $m_1 \geq \nu$). Additionally, $m_1 - m_2 = \nu - \mu - 1$.*

Note that the potential $V(\bar{x})$ has no (local) maxima or minima due to its harmonicity.

REMARK. The remaining (more difficult) case $\sum_{i=1}^{\mu} \zeta_i + \sum_{j=\mu+1}^l \zeta_j = 0$ is treated in [8].

REMARK. The above theorem has a generalization to any \mathbb{R}^n ($n \geq 3$) with m_1 being the number of the critical points of index 1 and m_2 being the number of the critical points of index $n - 1$.

DEFINITION 1.2. Configurations of charges with all non-degenerate critical points and $m_1 + m_2 = \mu + \nu - 1$ are said to be *minimal*; see [14, p. 292].

REMARK. Minimal configurations occur if one places, for example, all charges of the same sign on a straight line. On the other hand, it is easy to construct generic non-minimal configurations of charges; see [14].

REMARK. The major difficulty of this problem is that the *lower bound* on the number of critical points of V given by Morse theory is known to be not exact. Therefore, since we are interested in an effective *upper bound*, the Morse theory arguments do not provide an answer.

The question about the maximum (if it exists) of the number of points of equilibrium of a non-degenerate configuration of charges in \mathbb{R}^3 was posed in [14, p. 293]. In fact, J. C. Maxwell in [12, Section 113] made an explicit claim answering exactly this question.

CONJECTURE 1.3 ([12], see also the appendix below). The total number of points of equilibrium (all assumed to be non-degenerate) of any configuration with l charges in \mathbb{R}^3 never exceeds $(l - 1)^2$.

REMARK. In particular, there are at most four points of equilibrium for any configuration of three point charges according to Maxwell; see Figure 1.

REMARK. The above conjecture, buried among other material in [12], was apparently completely forgotten for over a 130 years and by chance (re)discovered by one of the authors.

Before formulating our results and conjectures let us first generalize the set-up. In the notation of Theorem 1.1 consider the family of potentials depending on a (now not necessarily integer) parameter $\alpha > 0$ and given by

$$V_{\alpha}(\bar{x}) = \left(\frac{\zeta_1}{\rho_1^{\alpha}} + \dots + \frac{\zeta_{\mu}}{\rho_{\mu}^{\alpha}} \right) + \left(\frac{\zeta_{\mu+1}}{\rho_{\mu+1}^{\alpha}} + \dots + \frac{\zeta_l}{\rho_l^{\alpha}} \right), \tag{1.2}$$

where $\rho_i = r_i^2$ for $i = 1, \dots, l$. (The choice of ρ_i instead of r_i is motivated by convenience of algebraic manipulations.) The limit case of $\alpha = 0$ corresponds to

$$V_0(\bar{x}) = \sum \zeta_i \log \rho_i,$$

that is, the logarithmic potential considered by Gauss.

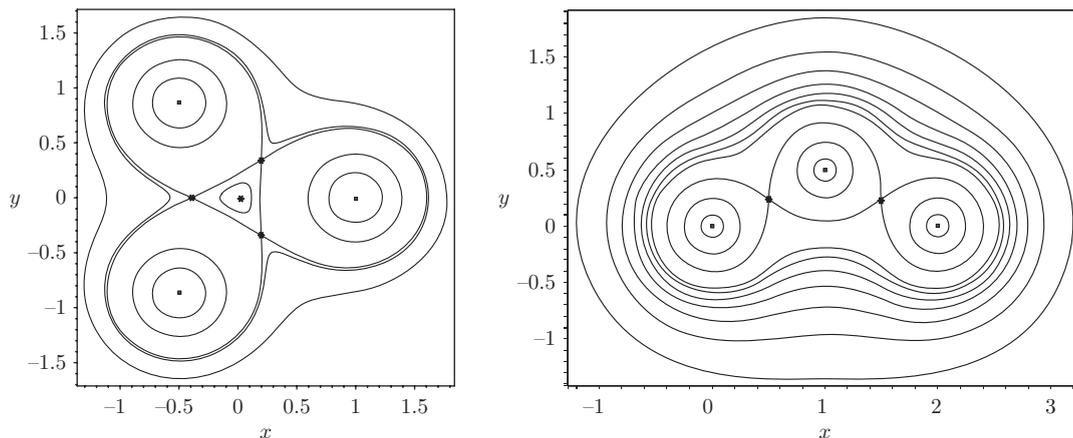


FIGURE 1. Configurations with two and with four critical points.

NOTATION 1.4. Denote by $N_l(n, \alpha)$ the maximal number of the critical points of the potential (1.2) where the maximum is taken over all non-degenerate configurations with l variable point charges, that is, over all possible values and locations of l point charges forming a non-degenerate configuration.

Our first result is the following uniform (that is, independent of n and α) upper bound.

THEOREM 1.5. (a) For any $\alpha \geq 0$ and any positive integer n one has

$$N_l(n, \alpha) \leq 4^{l^2} (3l)^{2l}. \tag{1.3}$$

(b) For $l = 3$ one has a significantly improved upper bound

$$N_3(n, \alpha) \leq 12.$$

REMARK. Note that the right-hand side of the formula (1.3) gives, even for $l = 3$, the horrible upper bound 139,314,069,504. On the other hand, computer experiments suggest that Maxwell was right and that for any three charges there are at most four (and not twelve) critical points of the potential (1.2); see Figure 1.

REMARK. Figure 1 shows the level curves of the restrictions of the potential of three positive charges to the plane they span in two essentially different cases (conjecturally, the only ones). The graph on the left has three saddles and one local minimum and the graph on the right has just two saddle points.

1.1. Voronoi diagrams and the main conjecture

Theorem 1.7 below determines the number of critical points of the function V_α for large α in terms of the combinatorial properties of the configuration of the charges. To describe it we need to introduce several notions.

NOTATION. By a (classical) *Voronoi diagram* of a configuration of pairwise distinct points (called sites) in the Euclidean space \mathbb{R}^n we understand the partition of \mathbb{R}^n into convex cells according to the distance to the nearest site; see, for example, [3, 17].

REMARK 1.6. The first known application of Voronoi diagrams can be traced back to Aristotle's *De Caelo* where Aristotle asked how a dog faced with the choice of two equally tempting meals could rationally choose between the two. These ideas were later developed by the well-known French philosopher and physicist Jean Buridan (1300–1356) who sowed the seeds of religious scepticism in Europe. Buridan allowed that the will could delay the choice in order to more fully assess the possible outcomes of the choice. Later writers satirized this view in terms of an ass who, confronted by two equally desirable and accessible bales of hay, must necessarily starve while pondering a decision. Apparently the Roman Catholic Church found unrecoverable errors in Buridan's arguments since about hundred and twenty years after his death a posthumous campaign by Okhamists succeeded in having Buridan's writings placed on the *Index Librorum Prohibitorum* (List of Forbidden Books) from 1474–1481.

A *Voronoi cell* S of the Voronoi diagram consists of all points having exactly the same set of nearest sites. The set of all nearest sites of a given Voronoi cell S is denoted by $\mathcal{NS}(S)$. Using the terminology of [3, 17], we see that the $\mathcal{NS}(S)$ is the set of vertices of the cell of the Delaunay triangulation corresponding to the cell S . One can see that each Voronoi cell is the interior of a convex polyhedron, probably of positive codimension. This is a slight generalization of traditional terminology, which considers the Voronoi cells of the highest dimension only.

A Voronoi cell of the Voronoi diagram of a configuration of sites is said to be *effective* if it intersects the convex hull of $\mathcal{NS}(S)$.

If we have an additional affine subspace $L \subset \mathbb{R}^n$, we say that a Voronoi cell S of the Voronoi diagram of a configuration of charges in \mathbb{R}^n is *effective with respect to L* if S intersects the (closed) convex hull of the orthogonal projection of $\mathcal{NS}(S)$ onto L .

A configuration of points is said to be *generic* if any Voronoi cell S of its Voronoi diagram of any codimension k has exactly $k + 1$ nearest sites and does not intersect the boundary of the convex hull of $\mathcal{NS}(S)$. Notice that one can show that if the cell S does not intersect the interior of the convex hull of $\mathcal{NS}(S)$ then it consists of *topological regular points* of the function 'distance to the nearest site' in the terminology of [19].

A subspace L intersects a Voronoi diagram *generically* if it intersects all its Voronoi cells transversally, any Voronoi cell S of codimension k intersecting L has exactly $k + 1$ nearest sites, and S does not intersect the boundary of the convex hull of the orthogonal projection of $\mathcal{NS}(S)$ onto L .

The *combinatorial complexity* of a given configuration of points is the total number of cells of all dimensions in its Voronoi diagram, and the *effective combinatorial complexity* of a given configuration of points is the total number of effective cells of all dimensions in its Voronoi diagram.

EXAMPLE. A Voronoi diagram of three non-collinear points A, B, C on the plane consists of seven Voronoi cells:

- (i) three two-dimensional cells S_A, S_B, S_C with $\mathcal{NS}(S)$ consisting of one point;
- (ii) three one-dimensional cells S_{AB}, S_{AC}, S_{BC} with $\mathcal{NS}(S)$ consisting of two points; for example, S_{AB} is a part of the perpendicular bisector of the segment $[A, B]$;
- (iii) one zero-dimensional cell S_{ABC} with $\mathcal{NS}(S)$ consisting of all three points; this is the point equidistant from all three points.

There are two types of generic configurations of three points; see Figure 3. The first type is of an acute triangle ΔABC and then all Voronoi cells are effective. The second type is of an obtuse triangle ΔABC and then (for the obtuse angle A) the Voronoi cells S_{BC} and S_{ABC} are not effective.

The case of the right-angled triangle ΔABC is non-generic: the cell S_{ABC} , though effective, lies on the boundary of the triangle.

The following result motivates our main conjecture, Conjecture 1.8 below.

THEOREM 1.7. (a) *For any generic configuration of point charges of the same sign there exists $\alpha_0 > 0$ such that for any $\alpha \geq \alpha_0$ the critical points of the potential $V_\alpha(\bar{x})$ are in one-to-one correspondence with effective cells of positive codimension in the Voronoi diagram of the considered configuration. The Morse index of each critical point coincides with the dimension of the corresponding Voronoi cell.*

(b) *Suppose that an affine subspace L intersects generically the Voronoi diagram of a given configuration of point charges of the same sign. Then there exists $\alpha_0 > 0$ (depending on the configuration and L) such that for any $\alpha \geq \alpha_0$ the critical points of the restriction of the potential $V_\alpha(\bar{x})$ to L are in one-to-one correspondence with cells that are effective with respect to L and are of positive codimension in the Voronoi diagram of the considered configuration. The Morse index of each critical point coincides with the dimension of the intersection of the corresponding Voronoi cell with L .*

More exactly, we prove below that critical points of the potential corresponding to an effective cell S lie on the distance $O(\alpha^{-1})$ from the point of intersection of S and of the convex hull of $\mathcal{NS}(S)$. Note that the genericity assumption in the theorem is essential: it is easy to see that in the case of the right-angled triangle the zero-dimensional cell, though effective, does not correspond to a critical point of a potential.

Finally, our computer experiments in one- and two-dimensional cases led us to the following optimistic conjecture.

CONJECTURE 1.8. (a) *For any generic configuration of unit point charges and any $\alpha \geq \frac{1}{2}$ one has*

$$a_\alpha^j \leq \sharp^j, \tag{1.4}$$

where a_α^j is the number of the critical points of index j of the potential $V_\alpha(\bar{x})$ and \sharp^j is the number of all effective Voronoi cells of dimension j in the Voronoi diagram of the considered configuration.

(b) *For any affine subspace L generically intersecting the Voronoi diagram of a given configuration of unit point charges one has*

$$a_{\alpha,L}^j \leq \sharp_L^j, \tag{1.5}$$

where $a_{\alpha,L}^j$ is the number of the critical points of index j of the potential $V_\alpha(\bar{x})$ restricted to L and \sharp_L^j is the number of all Voronoi cells with $\dim(S \cap L) = j$ effective with respect to L in the Voronoi diagram of the considered configuration.

We will refer to the inequality (1.4) as the *Maxwell inequality*, and to (1.5) as the *relative Maxwell inequality*.

REMARK. Theorem 1.7 and Conjecture 1.8 were inspired by two observations. On one hand, the critical points of $V_\alpha(\bar{x})$ and of $V_\alpha^{-1/\alpha}(\bar{x})$ are the same (one needs here positiveness of charges). However, the limit of $V_\alpha^{-1/\alpha}(\bar{x})$ when $\alpha \rightarrow \infty$ can be easily computed. Namely, one can easily show that

$$\lim_{\alpha \rightarrow \infty} V_\alpha^{-1/\alpha}(\bar{x}) = V_\infty(\bar{x}) = \min_{i=1,\dots,l} \rho_i(\bar{x}).$$

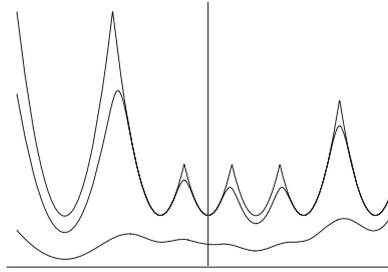


FIGURE 2. The potential $V_\alpha^{-1/\alpha}(x)$ converges to $V_\infty(x)$ as $\alpha \rightarrow \infty$.

Indeed, denoting $\rho(x) = \min_{i=1,\dots,l} \rho_i(\bar{x})$, we get

$$\begin{aligned} \lim_{\alpha \rightarrow \infty} V_\alpha^{-1/\alpha}(\bar{x}) &= \rho(\bar{x}) \lim_{\alpha \rightarrow \infty} \left[\sum_i \zeta_i \left(\frac{\rho(\bar{x})}{\rho_i(\bar{x})} \right)^\alpha \right]^{-1/\alpha} \\ &= \rho(\bar{x}) \lim_{\alpha \rightarrow \infty} (C_1 + C_2 e^{O(-\alpha)})^{-1/\alpha} = \rho(\bar{x}) = V_\infty(\bar{x}). \end{aligned}$$

This limiting function is only piecewise smooth. In the planar case in [19] an analogue of the standard Morse theory for this function was constructed. Generalizing this approach, we define critical points of $V_\infty(\bar{x})$ and their Morse indices. Moreover, it turns out that for generic configurations every critical point of $V_\infty(\bar{x})$ lies on a separate effective cell of the Voronoi diagram whose dimension equals the Morse index of that critical point; see § 2.4.3. Theorem 1.7 above claims that for sufficiently large α the situation is the same, except that the critical point does not lie exactly on the corresponding Voronoi cell (in fact, it lies an $O(\alpha^{-1})$ distance from this Voronoi cell; see Lemmas 2.30 and 2.35). On the other hand, computer experiments show that the largest number of critical points (if one fixes the positions of charges) occurs when $\alpha \rightarrow \infty$; see Figure 2.

Even the special case of Conjecture 1.8 when L is one-dimensional is of interest and still open. Its slightly stronger version supported by extensive numerical evidence can be reformulated as follows.

CONJECTURE 1.9. Consider an l -tuple of points $(x_1, y_1), \dots, (x_l, y_l)$ in \mathbb{R}^2 . Then for any values of charges $(\zeta_1, \dots, \zeta_l)$ the function $V_\alpha^*(x)$ in (one real) variable x given by

$$V_\alpha^*(x) = \sum_{i=1}^l \frac{\zeta_i}{((x - x_i)^2 + y_i^2)^\alpha} \tag{1.6}$$

has at most $(2l - 1)$ real critical points, provided one assumes that $\alpha \geq \frac{1}{2}$.

REMARK. In the simplest possible case $\alpha = 1$, Conjecture 1.9 is equivalent to showing that real polynomials of degree $(4l - 3)$ of a certain form have at most $(2l - 1)$ real zeros.

1.1.1. *Complexity of the Voronoi diagram and Maxwell's conjecture.* In the classical planar case one can show that the total number of cells of positive codimension of the Voronoi diagram of any l sites on the plane is at most $5l - 11$ and this bound is exact.

Since $(l - 1)^2$ is larger than the conjectural exact upper bound $5l - 11$ for all $l > 5$ and coincides with $5l - 11$ for $l = 3, 4$, we conclude that Conjecture 1.8 implies a stronger form of Maxwell's conjecture for any l positive charges on the plane and any $\alpha \geq \frac{1}{2}$.

For $n > 2$ the worst-case complexity $\Gamma(l, n)$ of the classical Voronoi diagram of an l -tuple of points in \mathbb{R}^n is $\Theta(l^{\lfloor n/2+1 \rfloor})$; see [3]. Namely, there exist positive constants $A < B$ such that $Al^{\lfloor n/2+1 \rfloor} < \Gamma(l, n) < Bl^{\lfloor n/2+1 \rfloor}$. Moreover, the Upper Bound Conjecture of the convex polytopes theory proved by McMullen implies that the number of Voronoi cells of dimension k of a Voronoi diagram of l charges in \mathbb{R}^n does not exceed the number of $(n - k)$ -dimensional faces in the $(n + 1)$ -dimensional cyclic polytope with l vertices; see [5, 13]. This bound is exact, that is, is achieved for some configurations; see [18].

In \mathbb{R}^3 this means that the number of 0-dimensional Voronoi cells of the Voronoi diagram of l points is at most $\frac{1}{2}l(l - 3)$, the number of 1-dimensional Voronoi cells is at most $l(l - 3)$, and the number of 2-dimensional Voronoi cells is at most $\frac{1}{2}l(l - 1)$.

We were unable to find a similar result about the number of *effective* cells of a Voronoi diagram. However, already for a regular tetrahedron the number of effective cells is 11, which is greater than Maxwell's bound 9. Thus a stronger version of Maxwell's conjecture in \mathbb{R}^3 fails: the number of critical points of V_α can be bigger than $(l - 1)^2$ for α sufficiently large.

However Maxwell's original conjecture miraculously agrees with Maxwell inequalities (1.4) and we obtain the following conditional statement.

THEOREM 1.10. *Conjecture 1.8 implies the validity of the original Maxwell conjecture for any configuration of positive charges in \mathbb{R}^3 in the standard 3-dimensional Newton potential, that is, $\alpha = \frac{1}{2}$.*

Existing literature and outline of the paper. Logarithmic potentials in \mathbb{R}^3 similar to (1.2) (that is, the case of the electrostatic force proportional to the inverse of the distance) were studied in a number of papers of J. L. Walsh; see [15] and references therein. In this case it is possible to generalize the classical Gauss–Lucas theorem and some results on Jensen's circles for polynomials in one complex variable to real vector spaces of higher dimension.

Some interesting examples of electrostatic potentials whose critical points form curves were considered in [6]. The question of whether degenerate electrostatic potential defined by a finite number of charges can have an analytic arc of critical points was stated in [14, p. 294]. Finally, the results about instability of critical points for more general potentials and dynamical systems are obtained in [9]. Instability in our context follows from subharmonicity of the considered potential and had already been mentioned in [12, Section 116] under the name Earnshaw's theorem.

The structure of the paper is as follows. In Section 2 we prove the above results. Section 3 contains further remarks and open problems related to the topic. Finally, in the appendix we reproduce the original Section 113 of [12] where Maxwell presents the arguments of Morse theory (developed at least 50 years later), and names the ranks of the 1st and the 2nd homology groups of domains in \mathbb{R}^3 in the language of (apparently existing) topology of the 1870s to formulate his claim.

Acknowledgements. The authors are sincerely grateful to A. Eremenko, A. Fryntov, D. Khavinson, H. Shapiro, M. Shapiro and A. Vainshtein for valuable discussions and references, and to the anonymous referee for very careful and helpful comments. B. Shapiro wants to acknowledge the hospitality of the Department of Mathematics, Purdue University during his visit in the spring of 2003.

2. Proofs

We start this section with a discussion of the non-degeneracy requirement and the (co)dimension of the affine span of a configuration of point charges.

2.1. Relation between the number of charges and dimension

Consider a non-degenerate configuration of $l = \mu + \nu$ point charges in \mathbb{R}^n , and let $L \subseteq \mathbb{R}^n$ be the affine subspace spanned by the points where the charges are located. Evidently, we have $\dim L \leq l - 1$.

THEOREM 2.1. *If all critical points of the potential V_α are isolated, then either all critical points belong to L or $n \leq l - 1$.*

Let us first show that it is enough to consider the cases $n \leq l$ only.

LEMMA 2.2. *If a configuration of charges in \mathbb{R}^n has only isolated critical points then either all its critical points belong to L or L is a hyperplane in \mathbb{R}^n .*

Proof. Indeed, assume that there is a critical point outside L and $\text{codim } L > 1$. Then the whole orbit of this point under the action of the group of rotations of \mathbb{R}^n preserving L consists of critical points (since this action preserves the potential). \square

To complete the proof one has to exclude the case $n = l$. We show that if $\dim L = l - 1$ then all critical points of the potential are in L .

LEMMA 2.3. *If one can find a hyperplane H in L separating positive charges from the negative ones, then the potential of the configuration has no critical points outside L .*

Proof. Indeed, let $x \notin L$ be any point outside L , and let H_x be any hyperplane containing both x and H and transverse to L . Let n be a vector normal to H_x at x . The signs of scalar products of the gradients of the potentials of each charge with n are the same, so x cannot be an equilibrium point. \square

Physically, the statement is obvious: all negative charges lie one side of H_x while all positive charges lie on its other side. A positive test charge lying on H_x is pushed away from positive charges and towards the negative ones. By the above assumptions, the resulting force is non-vanishing and directed towards the half-space containing the negative charges.

COROLLARY 2.4. *The potential of any configuration of positive charges has no critical points outside L .*

COROLLARY 2.5. *Any configuration with l point charges such that $\dim L = l - 1$ has no critical points of the potential outside L .*

Proof. These points should form a non-degenerate simplex, and any subset of vertices of a simplex can be separated from the rest of the vertices by a hyperplane, so the claim follows from the previous lemma. \square

REMARK. As one can see from the proof, the set of critical points of a configuration is a union of spheres with centers in L and of dimension equal to $\text{codim } L$. As n grows, the dimension of spheres is the only parameter that changes, so the case $\text{codim } L = 1$ is the most general one.

We conclude that in any case it is enough to consider the case $n \leq l - 1$.

2.2. Proof of Theorem 1.5(a)

The proof is an application of the theory of fewnomials developed by A. G. Khovanskii in [7]. A serious drawback of this theory is that the obtained estimates, though effective, are usually highly excessive. Applying the methods, rather than the results, of this theory one might get a much better estimate which we illustrate while proving part (b) of Theorem 1.5.

We start with the following result from [7, §1.2].

THEOREM 2.6 (Khovanskii). *Consider a system of m quasipolynomial equations*

$$P_1(\bar{u}, \bar{w}(\bar{u})) = \dots = P_m(u, \bar{w}(\bar{u})) = 0, \quad \text{where } \bar{u} = (u_1, \dots, u_m),$$

where each P_i is a real polynomial of degree d_i in $(m + k)$ variables $(u_1, \dots, u_m, w_1, \dots, w_k)$ and

$$w_j = \exp\langle \bar{a}_j, \bar{u} \rangle, \quad \bar{a}_j = (a_j^1, \dots, a_j^m) \in \mathbb{R}^m, \quad \text{for } j = 1, \dots, k.$$

Then the number of real isolated solutions of this system does not exceed

$$d_1 \dots d_m (d_1 + \dots + d_m + 1)^k 2^{k(k-1)/2}.$$

The estimate of Theorem 1.5(a) will follow from a presentation of the critical points of a configuration of point charges as solutions to an appropriate system of quasipolynomial equations; see below.

2.2.1. Constructing a quasipolynomial system. Consider a configuration with l point charges in \mathbb{R}^n . Denote by $\bar{x} = (x_1, \dots, x_n)$ the coordinates of a critical point and denote by (c_1^i, \dots, c_n^i) , for $i = 1, \dots, l$, the coordinates of the i th charge. We assume that the 1st charge is placed at the origin, that is, that $c_1^1 = \dots = c_n^1 = 0$.

The first l equations of our system define the variables $\bar{\rho} = (\rho_1, \dots, \rho_l)$ as the squares of distances between the variable point \bar{x} and the charges. They can be rewritten as

$$P_1(\bar{x}, \bar{\rho}) = \dots = P_l(\bar{x}, \bar{\rho}) = 0, \tag{2.1}$$

where

$$P_1(\bar{x}, \bar{\rho}) = \sum_{j=1}^n x_j^2 - \rho_1, \quad P_i(\bar{x}, \bar{\rho}) = \rho_1 - \rho_i + \sum_{j=1}^n c_j^i (-2x_j + c_j^i), \quad \text{for } i = 2, \dots, l. \tag{2.2}$$

The second group of equations expresses the fact that the point $\bar{x} = (x_1, \dots, x_n)$ is the critical point of the potential $V_\alpha(\bar{x}) = \sum \zeta_i \rho_i^{-\alpha}$. Namely,

$$\frac{\partial}{\partial x_j} V_\alpha(\bar{x}) = \sum_{i=1}^l \zeta_i \frac{\partial}{\partial x_j} \rho_i^{-\alpha} = -2\alpha \sum_{i=1}^l \zeta_i h_i (x_j - c_j^i) = P_{l+j}(\bar{x}, \bar{h}), \quad \text{for } j = 1, \dots, n,$$

where we denote

$$h_i = \rho_i^{-\alpha-1}, \quad \text{for } i = 1, \dots, l, \quad \text{and} \quad \bar{h} = (h_1, \dots, h_l). \tag{2.3}$$

Introducing new variables $s_i = \log \rho_i$ we get the system

$$P_1(\bar{x}, \bar{s}, \bar{\rho}, \bar{h}) = \dots = P_{n+l}(\bar{x}, \bar{s}, \bar{\rho}, \bar{h}) = 0,$$

of $(l + n)$ quasipolynomial equations in $(l + n)$ variables $(\bar{x}, \bar{s}, \bar{\rho}, \bar{h})$ with

$$\rho_i = \exp(s_i), \quad h_i = \exp(-(\alpha + 1)s_i), \quad \text{for } i = 1, \dots, l.$$

This system has the type described in Theorem 2.6, with $m = n + l$, $k = 2l$,

$$\deg P_1 = \deg P_{l+1} = \dots = \deg P_{n+k} = 2, \quad \text{and} \quad \deg P_2 = \dots = \deg P_l = 1.$$

By Proposition 2.1, one has $n \leq l - 1$ which implies the required estimate:

$$N_l(n, \alpha) \leq N_l(l - 1, \alpha) \leq 4^{l^2} 9^l l^{2l} = 4^{l^2} (3l)^{2l}.$$

For example, for $l = 3$ one gets $N_3(n, \alpha) \leq 139, 314, 069, 504$.

2.3. *Proof of Theorem 1.5(b)*

As we mentioned above, one can do much better by applying the fewnomials method rather than results, and here we demonstrate this in the case of three charges. By Proposition 2.1 we can restrict our consideration to the case of $\mathbb{R}^{l-1} = \mathbb{R}^2$ (however the exponent α in the potential V_α can be an arbitrary positive number).

The scheme of this rather long proof is as follows. First, we pass to more convenient coordinates. In these new coordinates the equilibrium points correspond to the intersection points of two explicitly written planar curves γ_1 and γ_2 in the positive quadrant \mathbb{R}_+^2 of the real plane. Both curves are separating solutions of Pfaffian forms. Two consecutive applications of the Rolle–Khovanskii theorem result in two real polynomials R and Q such that the required upper bound can be given in terms of the number of their common zeros in \mathbb{R}_+^2 . The latter is bounded by the Bernstein–Kushnirenko bound and then decreased by the number of common roots of R and Q known to be outside of \mathbb{R}_+^2 .

2.3.1. *Changing variables and getting a system of equations.* To emphasize that the methods given below can be generalized we make the change of variables in the situation of l charges in \mathbb{R}^n (we assume, as before, that $n \leq l - 1$). This, as a by-product, produces another proof of Theorem 1.5 with a somewhat better upper bound.

As above, we assume that the charges ζ_i are located at (c_1^i, \dots, c_n^i) , for $i = 1, \dots, l$, and assume for simplicity that $c_1^1 = \dots = c_n^1 = 0$. We consider the potential

$$V_\alpha(x_1, \dots, x_n) = \sum_{i=1}^l \zeta_i \rho_i^{-\alpha}, \quad \text{where } \rho_i = \sum_{j=1}^n (x_j - c_j^i)^2 \text{ for } i = 1, \dots, l.$$

The system of equations defining the critical points of $V_\alpha(\bar{x})$ is

$$\frac{\partial V_\alpha(\bar{x})}{\partial x_j} = 0 \quad \text{for } j = 1, \dots, n, \quad \text{where } \frac{\partial V_\alpha(\bar{x})}{\partial x_j} = -2\alpha \sum_{i=1}^l \zeta_i \rho_i^{-\alpha-1} (x_j - c_j^i).$$

Introducing $h_i = \rho_i^{-\alpha-1}$ one can solve each equation of this system and express x_j in terms of h_i :

$$x_j = \frac{\sigma_j}{\sigma}, \quad \text{where } \sigma = \sum_{i=1}^l \zeta_i h_i \text{ and } \sigma_j = \sum_{i=1}^l \zeta_i h_i c_j^i \tag{2.4}$$

are homogeneous linear functions of h_i . The equilibrium points correspond to the solutions of the following system of equations obtained from the definition of h_i by substitution of σ_j/σ instead of x_j :

$$h_i^{-1/(\alpha+1)} = \frac{\xi_i}{\sigma^2} \quad \text{for } i = 1, \dots, l, \quad \text{where } \xi_i = \sum_{j=1}^n (\sigma_j - c_j^i \sigma)^2. \tag{2.5}$$

This system has the following remarkable properties.

PROPOSITION 2.7. (a) Any solution of $\sigma = \xi_1 = 0$ is also a zero of all the ξ_i .

(b) Each ξ_i is a strictly positive real quadratic polynomial that is homogeneous in h and independent of h_i .

Proof. Indeed, $\xi_1 - \xi_i = \sigma \sum_{j=1}^n c_j^i (2\sigma_j - \sigma c_j^i)$.

The second statement is obvious except for the independence of h_i , which can be proved by direct computation. \square

REMARK. Note that the above system can be represented as a system of quasipolynomials as in Theorem 2.6. Namely, the equations in (2.5) are polynomials in $h_i, h_i^{1/(\alpha+1)}$. Introducing $s_i = \log(\sigma_i/\sigma_1)$ one can apply Theorem 2.6. After several small tricks — dehomogenization of the system, introduction of a new variable $z = \xi_1$ and noticing that the expression for $\xi_1 - \xi_i$ then becomes linear — we obtain an upper bound $2 \cdot 4^{l^2} (2l + 3)^{2l}$ on the number of equilibrium points of a system of l charges. For $l > 3$ this bound is somewhat better than the bound (1.3).

Now, let us use the previous construction for $l = 3$ and $n = 2$. To save subscripts, we denote coordinates in \mathbb{R}^2 by (x, y) instead of (x_1, x_2) . Without loss of generality we can assume that the three charges with the values ζ_1, ζ_2 and 1 are located at $(0, 0), (1, 0)$ and (a, b) respectively.

Expressions (2.4) are homogeneous in h_j , so we introduce the non-homogeneous variables f and g as follows:

$$f = \frac{h_2}{h_1} = \left(\frac{\rho_1}{\rho_2}\right)^{\alpha+1} \quad \text{and} \quad g = \frac{h_3}{h_1} = \left(\frac{\rho_1}{\rho_3}\right)^{\alpha+1}. \tag{2.6}$$

Then equations (2.4) become

$$x = \frac{ag + \zeta_2 f}{\zeta_1 + \zeta_2 f + g}, \quad y = \frac{bg}{\zeta_1 + \zeta_2 f + g}. \tag{2.7}$$

The system (2.5) reduces to the following two equations describing two curves γ_1 and γ_2 in the positive quadrant \mathbb{R}_+^2 of the (f, g) -plane:

$$\gamma_1 = \left\{ f^{1/(\alpha+1)} \xi_2 \xi_1^{-1} = 1 \right\}, \quad \gamma_2 = \left\{ g^{-1/(\alpha+1)} \xi_2 = f^{-1/(\alpha+1)} \xi_3 \right\}. \tag{2.8}$$

Here

$$\begin{aligned} \xi_1 &= (ag + \zeta_2 f)^2 + b^2 g^2, \\ \xi_2 &= ((a - 1)g - \zeta_1)^2 + b^2 g^2, \\ \xi_3 &= ((a - 1)\zeta_2 f + a\zeta_1)^2 + b^2 (\zeta_2 f + \zeta_1)^2. \end{aligned}$$

The following facts about ξ_i follow from Proposition 2.7.

PROPOSITION 2.8. *The following hold:*

- (i) ξ_2 depends only on g , and ξ_3 depends only on f ;
- (ii) ξ_2 and ξ_3 are strictly positive quadratic polynomials;
- (iii) ξ_1 is a positive definite homogeneous quadratic form;
- (iv) $\xi_1 = \xi_2 = \xi_3 = 0$ have two complex solutions.

The goal of all subsequent computations is to give an upper bound on the number N of the intersection points between γ_1 and γ_2 in \mathbb{R}_+^2 . We are able to obtain the following estimate proved below.

PROPOSITION 2.9. *The number of intersection points of γ_1 and γ_2 lying in \mathbb{R}_+^2 is at most 12.*

We note that any intersection point of γ_1 and γ_2 lying in the positive quadrant $\mathbb{R}_+^2 = \{f > 0, g > 0\}$ corresponds, via (2.7), to the unique critical point of $V_\alpha(x, y)$, so the estimate of Theorem 1.5(b) immediately follows.

2.3.2. *The Rolle–Khovanskii theorem.* Before we move further let us recall the \mathbb{R}^2 -version of a generalization of Rolle’s theorem due to Khovanskii.

Suppose that we are given a smooth differential 1-form ω defined in a domain $D \subset \mathbb{R}^2$. Let $\gamma \subset D$ be a (not necessarily connected) one-dimensional integral submanifold of ω .

DEFINITION 2.10. We say that γ is a *separating solution* of ω if

- (a) γ is the boundary of some (not necessarily connected) domain U ;
- (b) the coorientations of γ defined by ω and by U coincide (that is, ω is positive on the outer normal to the boundary of U).

Let γ_1 and γ_2 be two separating solutions of two 1-forms ω_1 and ω_2 respectively.

THEOREM 2.11 (see [7]). *We have*

$$\sharp(\gamma_1, \gamma_2) \leq \sharp(\gamma_1) + \flat(\gamma_1, \gamma_2),$$

where $\sharp(\gamma_1, \gamma_2)$ is the number of intersection points between γ_1 and γ_2 , $\sharp(\gamma_1)$ is the number of non-compact components of γ_1 , and $\flat(\gamma_1, \gamma_2)$ is the number of the points of contact of γ_1 and ω_2 , that is, the number of points of γ_1 such that $\omega_2(\dot{\gamma}_1) = 0$. (One can also characterize the latter points as the intersection points of γ_1 with an algebraic set $\{\omega_1 \wedge \omega_2 = 0\}$.)

2.3.3. *First application of the Rolle–Khovanskii theorem.* We apply Theorem 2.11 to the curves γ_1 and γ_2 defined in (2.8). These curves are integral curves in \mathbb{R}_+^2 of the one-forms η_1 and η_2 respectively, where

$$\eta_1 = \frac{df}{(\alpha + 1)f} + \frac{\xi'_2 dg}{\xi_2} - \frac{d\xi_1}{\xi_1}, \tag{2.9}$$

$$\eta_2 = \left(-\frac{1}{(\alpha + 1)f} + \frac{\xi'_3}{\xi_3} \right) df + \left(\frac{1}{(\alpha + 1)g} - \frac{\xi'_2}{\xi_2} \right) dg. \tag{2.10}$$

These forms are logarithmic differentials of the functions defining the curves: if we denote $F = (f/g)^{-1/(\alpha+1)} \xi_3 \xi_2^{-1}$, then $\gamma_2 = \{F = 1\}$ and $\eta_2 = d \log F$. Similarly, $\eta_1 = d \log G$, where $G = f^{1/(\alpha+1)} \xi_2 \xi_1^{-1}$.

In what follows, we assume that 1 is a regular value of F and G . This can always be achieved by a small perturbation of parameters and is enough for the proof of Theorem 1.5(b) by upper-continuity of the number of the non-degenerate critical points.

LEMMA 2.12. *The curves γ_1 and γ_2 are separating solutions of the polynomial forms η_1 and η_2 respectively.*

Proof. Indeed, γ_2 is a level curve of the function $F = (f/g)^{-1/(\alpha+1)} \xi_3 \xi_2^{-1}$, which is a smooth function on \mathbb{R}_+^2 . Thus, γ_2 coincides with the boundary of the domain $\partial\{F < 1\}$. Therefore the value of $\eta_2 = d(\log F)$ on the outer normal to $\{F < 1\}$ is non-negative, and is everywhere positive since 1 is not a critical value of F . This means that the coorientations of γ_2 as the boundary of $\{F < 1\}$ and as defined by the polynomial form η_2 coincide.

Similar arguments hold for γ_1 , and we conclude that γ_1 and γ_2 are separating solutions of the forms η_1 and η_2 . □

This enables application of Theorem 2.11 to the pair (γ_1, γ_2) and we obtain the following estimate.

PROPOSITION 2.13. *We have*

$$N \leq N_1 + N_2,$$

where N is the number of points in the intersection $\gamma_1 \cap \gamma_2 \cap \mathbb{R}_+^2$, N_1 is the number of the non-compact components of γ_2 in \mathbb{R}_+^2 , and N_2 is the number of points of intersection of γ_2 with the set $\Gamma = \{\eta_1 \wedge \eta_2 = 0\}$ in \mathbb{R}_+^2 .

LEMMA 2.14. $N_1 = 2$.

Proof. Indeed, if we denote $\beta = -1/(\alpha + 1)$, then the equation $f^{-1/(\alpha+1)}\xi_3 = g^{-1/(\alpha+1)}\xi_2$ can be rewritten as

$$\zeta_2^2 f^{\beta+2} + C_1 f^{\beta+1} + C_2 f^\beta = g^{\beta+2} + C_3 g^{\beta+1} + C_4 g^\beta,$$

with some constants C_i . Since $-1 < \beta < 0$ and $C_2, C_4 > 0$, the left-hand side tends monotonically to $+\infty$ as f tends to ∞ or to zero, and similarly for the right-hand side. Therefore as $g \gg 1$ or $0 < g \ll 1$, the right-hand side of the resulting equation in f has two positive solutions, one close to zero, and another close to infinity. For bounded g , with $g \in [\epsilon_1, \epsilon_1^{-1}]$, the solution $f = f(g)$ is bounded as well: $f \in [\epsilon_2, \epsilon_2^{-1}]$ for some $\epsilon_2 \ll \epsilon_1$. We conclude that the number of intersection points of γ_2 with the boundary of a large rectangle

$$\{\epsilon_1 \leq f \leq \epsilon_1^{-1}, \epsilon_2 \leq g \leq \epsilon_2^{-1}, 0 < \epsilon_1 \ll \epsilon_2 \ll 1\}$$

is 4. The number of non-compact components of γ_2 in \mathbb{R}_+^2 is equal to half of this number, that is, to 2. □

REMARK. By a slightly more careful analysis one can show that the asymptotical behavior of four ends of γ_2 is as follows: $g \sim \text{const} \cdot f$ as $f \rightarrow 0$ or ∞ , $g \sim \text{const} \cdot f^{-1-2\alpha}$ as $f \rightarrow \infty$, and $f \sim \text{const} \cdot g^{-1-2\alpha}$ as $g \rightarrow \infty$.

2.3.4. *Second application of the Rolle–Khovanskii theorem.* In order to estimate N_2 we apply Theorem 2.11 again. The set $\Gamma = \{\eta_1 \wedge \eta_2 = 0\}$ is a real algebraic curve given by the equation $Q = 0$, where

$$Q df \wedge dg = fg\xi_1\xi_2\xi_3 \cdot \eta_1 \wedge \eta_2 \tag{2.11}$$

is a polynomial in (f, g) . Applying Theorem 2.11 again we get the following estimate.

PROPOSITION 2.15. *We have*

$$N_2 \leq N_3 + N_4,$$

where (as above) N_2 is the number of points in $\{\gamma_2 \cap \Gamma \cap \mathbb{R}_+^2\}$, N_3 is the number of non-compact components of Γ in \mathbb{R}_+^2 , and N_4 is the number of points in $\Gamma \cap \{dQ \wedge \eta = 0\} \cap \mathbb{R}_+^2$.

NOTATION 2.16. For any polynomial S in two variables let $\mathcal{NP}(S)$ denote the Newton polygon of S . We denote the partial order on the set of plane polygons by inclusion by \prec , namely, $A \prec B$ means that a polygon A lies strictly inside a polygon B .

LEMMA 2.17. *The set Γ has no unbounded components in \mathbb{R}^2 . Moreover, Γ does not intersect the coordinate axes except at the origin, which is an isolated point of Γ .*

Proof. We start by identifying the essential part of Q .

LEMMA 2.18. *We have*

$$Q = \frac{-1 - 2\alpha}{(\alpha + 1)^2} \xi_1 \xi_2 \xi_3 + fgQ_1, \quad \text{where } Q_1 = \xi_2' \xi_3 \frac{\partial \xi_1}{\partial f} + \xi_2 \xi_3' \frac{\partial \xi_1}{\partial g} - \xi_2' \xi_3' \xi_1. \tag{2.12}$$

Proof. We defined Q in (2.11) via one-forms η_i defined in (2.9). Rewrite η_i as a linear combination of df and dg (recall that ξ_2 depends only on g , and ξ_3 depends only on f):

$$\begin{aligned} \eta_1 &= \left(\frac{1}{(\alpha+1)f} - \frac{(\xi_1)_f}{\xi_1} \right) df + \left(\frac{\xi'_2}{\xi_2} - \frac{(\xi_1)_g}{\xi_1} \right) dg, \\ \eta_2 &= \left(-\frac{1}{(\alpha+1)f} + \frac{\xi'_3}{\xi_3} \right) df + \left(\frac{1}{(\alpha+1)g} - \frac{\xi'_2}{\xi_2} \right) dg. \end{aligned}$$

Calculate the ratio $\eta_1 \wedge \eta_2 / df \wedge dg$:

$$\begin{aligned} &\left(\frac{1}{(\alpha+1)f} - \frac{(\xi_1)_f}{\xi_1} \right) \left(\frac{1}{(\alpha+1)g} - \frac{\xi'_2}{\xi_2} \right) - \left(\frac{\xi'_2}{\xi_2} - \frac{(\xi_1)_g}{\xi_1} \right) \left(-\frac{1}{(\alpha+1)f} + \frac{\xi'_3}{\xi_3} \right) \\ &= \left(\frac{1}{(\alpha+1)^2 fg} - \frac{(\xi_1)_f}{(\alpha+1)\xi_1 g} - \frac{\xi'_2}{(\alpha+1)f\xi_2} + \frac{(\xi_1)_f \xi'_2}{\xi_1 \xi_2} \right) \\ &\quad - \left(-\frac{\xi'_2}{(\alpha+1)f\xi_2} + \frac{(\xi_1)_g}{(\alpha+1)f\xi_1} + \frac{\xi'_2 \xi'_3}{\xi_2 \xi_3} - \frac{(\xi_1)_g \xi'_3}{\xi_1 \xi_3} \right). \end{aligned}$$

Observe that the third and fifth terms cancel. The function ξ_1 is quadratic and homogeneous in f and g . Therefore the second and sixth terms can be rewritten using Euler's identity $2\xi_1 = f(\xi_1)_f + g(\xi_1)_g$ as

$$-\frac{(\xi_1)_f}{(\alpha+1)\xi_1 g} - \frac{(\xi_1)_g}{(\alpha+1)f\xi_1} = -\frac{f(\xi_1)_f + g(\xi_1)_g}{(\alpha+1)f g \xi_1} = -\frac{2}{(\alpha+1)fg}.$$

Therefore the ratio $\eta_1 \wedge \eta_2 / df \wedge dg$ is equal to

$$\begin{aligned} &\frac{1}{(\alpha+1)^2 fg} - \frac{2}{(\alpha+1)fg} + \frac{(\xi_1)_f \xi'_2}{\xi_1 \xi_2} + \frac{(\xi_1)_g \xi'_3}{\xi_1 \xi_3} - \frac{\xi'_2 \xi'_3}{\xi_2 \xi_3} \\ &= \frac{1}{fg\xi_1 \xi_2 \xi_3} \left(\frac{-1-2\alpha}{(1+\alpha)^2} \xi_1 \xi_2 \xi_3 + fg [(\xi_1)_f \xi'_2 \xi_3 + (\xi_1)_g \xi'_3 \xi_2 - \xi_1 \xi'_2 \xi'_3] \right), \end{aligned}$$

which is what is needed. □

One can easily check that

$$\frac{\partial^3 Q_1}{\partial f^3} = \frac{\partial^3 Q_1}{\partial g^3} = \frac{\partial^4 Q_1}{\partial f^2 \partial g^2} \equiv 0.$$

Indeed, Q_1 is a polynomial of degree at most 4. However, the terms of the highest degree 4 are

$$2c_2 c_3 f^2 g (\xi_1)_f + 2c_2 c_3 f g^2 (\xi_1)_g - 4c_2 c_3 f g \xi_1 = 2c_2 c_3 f g (f(\xi_1)_f + g(\xi_1)_g - 2\xi_1),$$

which is zero by Euler's identity for ξ_1 (here c_i denote the leading coefficients of ξ_i). Therefore Q_1 has degree 3 in f and g . Terms of degree 3 in f appear only in $(\xi_1)_f \xi'_2 \xi_3$ and $\xi_1 \xi'_2 \xi'_3$ and cancel each other. A similar result holds for g .

Also, $Q_1(0,0) = 0$ since ξ_1 is homogeneous in f and g of degree 2. This implies that the Newton polygon of the polynomial fgQ_1 lies strictly inside the Newton polygon of $\xi_1 \xi_2 \xi_3$:

$$\mathcal{NP}(fgQ_1) = \{3 \leq p+q \leq 5, 1 \leq p, q \leq 3\} \prec \mathcal{NP}(Q) = \{2 \leq p+q \leq 6, 0 \leq p, q \leq 4\}.$$

Therefore the number of unbounded components of Γ in \mathbb{R}_+^2 coincides with the number of unbounded components of the zero locus of $\xi_1 \xi_2 \xi_3$ in \mathbb{R}_+^2 , the latter being equal to zero.

Another proof can be obtained by parameterizing the unbounded components of Γ near infinity and near the axes as $(f = Bt^{\epsilon_1} + \dots; g = At^{\epsilon_2} + \dots)$. From the shape of the Newton polygon of Q one can show that ϵ_1/ϵ_2 is either 0, 1 or ∞ . Therefore, B should be a root of ξ_3 , A/B should be a root of ξ_1 , or A should be a root of ξ_2 , respectively. Since neither of them has real roots, we conclude that Γ has no real unbounded components.

On the coordinate axes the polynomial Q equals $\xi_1\xi_2\xi_3$ and is therefore positive with the exception of the origin. The quadratic form of Q at the origin, being proportional to ξ_1 , is definite. Therefore the origin is an isolated zero of Q and, therefore, an isolated point of Γ . \square

COROLLARY 2.19. $N_3 = 0$.

2.3.5. *Estimating the number N_4 of points of contact between Γ and η_2 .* These points are the zeros of the polynomial form

$$Rdf \wedge dg = fg\xi_2\xi_3 \cdot dQ \wedge \eta_2.$$

Thus we have to estimate the number of solutions of the system $Q = R = 0$ in \mathbb{R}_+^2 .

We proceed as follows. Using the Bernstein–Kushnirenko theorem we find an upper bound on the number of solutions of $Q = R = 0$ in $(\mathbb{C}^*)^2$, and then reduce it by the number of solutions known to lie outside \mathbb{R}_+^2 .

The Bernstein–Kushnirenko upper bound for the number of common zeros of polynomials Q and R is expressed in terms of the mixed volume of their Newton polygons. In fact, in computation of this mixed volume we replace R by its difference with a suitable multiple of Q : this operation does not change common zeros of Q and R , but might significantly decrease the mixed volume of their Newton polygons.

Simple degree count shows that the Newton polygon of R is given by

$$\mathcal{NP}(R) = \{2 \leq p + q \leq 10, 0 \leq p, q \leq 6\}.$$

LEMMA 2.20. *There exists a polynomial $q = q(f, g)$ such that the Newton polygon of $\tilde{R} = R - qQ$ lies strictly inside the Newton polygon of R . In other words,*

$$\mathcal{NP}(\tilde{R}) \subseteq \{3 \leq p + q \leq 9, 1 \leq p, q \leq 5\}.$$

Proof. Our goal is to prove that all monomials lying on the boundary of $\mathcal{NP}(R)$ (further called boundary monomials) are equal to monomials lying on the boundary of a Newton polygon of some multiple of Q . We constantly use the fact that the boundary monomials of a product are equal to the boundary monomials of the product of boundary monomials of the factors.

First, let us replace R by a polynomial R_2 with the same Newton polygon and the same monomials on its boundary, but with a simpler definition. We have seen above that

$$\mathcal{NP}(fgQ_1) \prec \mathcal{NP}(Q) = \mathcal{NP}(\xi_1\xi_2\xi_3).$$

Denote $R_1df \wedge dg = fg\xi_2\xi_3 \cdot d(fgQ_1) \wedge \eta_2$. Computation of degrees shows that

$$\mathcal{NP}(R_1) \subseteq \{3 \leq p + q \leq 9, 1 \leq p, q \leq 5\} \prec \mathcal{NP}(R) = \{2 \leq p + q \leq 10, 0 \leq p, q \leq 6\}.$$

Therefore one can disregard R_1 and consider only $R_2 = R - R_1$, where

$$R_2df \wedge dg = fg\xi_2\xi_3 \cdot d(Q - fgQ_1) \wedge d\eta_2 = \text{const} \cdot fg\xi_2\xi_3 \cdot d(\xi_1\xi_2\xi_3) \wedge \eta_2.$$

Computing the product we see that

$$\begin{aligned} R_2 &= \text{const} \cdot d(\xi_1\xi_2\xi_3) \wedge \left[g\xi_2 \left(-\frac{\xi_3}{1+\alpha} + f\xi_3' \right) df + f\xi_3 \left(\frac{\xi_2}{1+\alpha} - g\xi_2' \right) dg \right] \\ &= \text{const} \cdot \xi_2\xi_3 \left\{ \xi_1 [2\xi_2\xi_3 + f\xi_2\xi_3' + g\xi_3\xi_2' - 3(1+\alpha)fg\xi_2'\xi_3'] - (1+\alpha)fgQ_1 \right\}. \end{aligned}$$

Simple computation using (2.8) shows that the Newton polygon of the first product in the figure brackets coincides with $\mathcal{NP}(Q) = \mathcal{NP}(\xi_1\xi_2\xi_3)$. The Newton polygon of the second product lies strictly inside of $\mathcal{NP}(Q)$, as was shown in (2.12). Therefore it does not affect boundary monomials and can be disregarded.

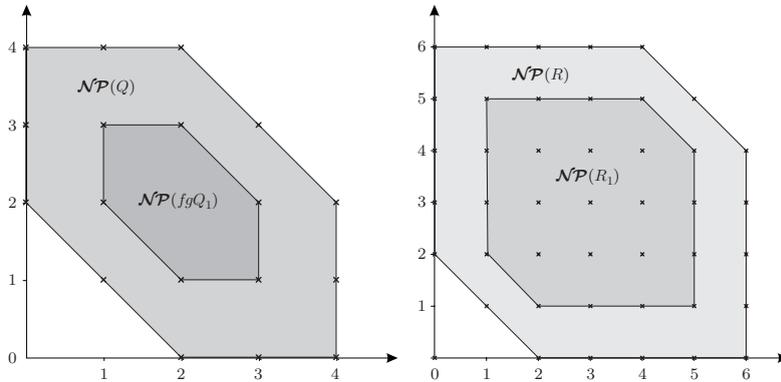


FIGURE 3. Relevant Newton polygons.

The remaining terms sum to $q\xi_1\xi_2\xi_3$, where we denote

$$q = f\xi_2\xi'_3 + g\xi_3\xi'_2 + 2\xi_2\xi_3 - 3(1 + \alpha)fg\xi'_2\xi'_3.$$

Up to a non-zero constant factor, its boundary monomials are the same as the boundary monomials of qQ : the polynomials Q and $\xi_1\xi_2\xi_3$ have proportional boundary monomials by (2.12).

Using these facts we conclude that for $\tilde{R} = R - \text{const} \cdot qQ$ one gets $\mathcal{NP}(\tilde{R}) \prec \mathcal{NP}(R)$. \square

2.3.6. *The Bernstein–Kushnirenko theorem.* Applying the well-known result of [2, 10] we know that the number of common zeros of Q and \tilde{R} in $(\mathbb{C}^*)^2$ does not exceed twice the mixed volume of $\mathcal{NP}(Q)$ and $\mathcal{NP}(\tilde{R})$.

Recall the definition of the mixed volume of two polygons. Let A and B be two planar convex polygons. It is common knowledge that the volume of their Minkowsky sum $\lambda A + \mu B$ is a homogeneous quadratic polynomial in (positive) λ and μ :

$$\text{Vol}(\lambda A + \mu B) = \text{Vol}(A)\lambda^2 + 2 \text{Vol}(A, B)\lambda\mu + \text{Vol}(B)\mu^2.$$

By definition the *mixed volume* of two polygons A and B is the coefficient $\text{Vol}(A, B)$.

Setting $\lambda = \mu = 1$, one gets

$$2 \text{Vol}(A, B) = \text{Vol}(A + B) - \text{Vol}(A) - \text{Vol}(B).$$

LEMMA 2.21. *There are at most 28 common zeros of $Q = R = 0$ in $(\mathbb{C}^*)^2$.*

Proof. A simple count gives $2 \text{Vol}(\mathcal{NP}(Q), \mathcal{NP}(\tilde{R})) = 28$. See Figure 3. \square

2.3.7. *Common zeros of Q and R outside \mathbb{R}_+^2 .* By Lemma 2.8 the quadratic polynomials ξ_1, ξ_2 and ξ_3 have two common zeros. Denote them by (f_1, g_1) and (f_2, g_2) . Evidently, they are not real (since, for example, ξ_3 is strictly positive on \mathbb{R}^2).

LEMMA 2.22. *The points (f_1, g_1) and (f_2, g_2) are solutions of the system $Q = R = 0$ of multiplicity at least 6 each.*

Proof. Since (f_1, g_1) and (f_2, g_2) are conjugate and the system $Q = R = 0$ is real, it suffices to consider one of these points, say (f_1, g_1) .

From (2.12) one can immediately see that $Q(f_1, g_1) = 0$. Indeed, each term has a factor annihilating at (f_1, g_1) (recall that all three ξ_i are zero at this point).

Moreover, differentiating Q_1 , one can see that

$$\frac{\partial Q_1}{\partial f}(f_1, g_1) = \frac{\partial Q_1}{\partial g}(f_1, g_1) = 0,$$

so (f_1, g_1) is a critical point of Q .

Recall that

$$R = dQ \wedge \left(g\xi_2 \left(-\frac{\xi_3}{1+\alpha} + f\xi_3' \right) df + f\xi_3 \left(\frac{\xi_2}{1+\alpha} - g\xi_2' \right) dg \right),$$

that is, R is the product of two polynomial forms each having a simple zero at (f_1, g_1) . Therefore this point is necessarily a critical point of R as well.

Moreover, the 1-jet of η_2 at (f_1, g_1) equals

$$j^1\eta_2 = f_1g_1\xi_2'(g_1)\xi_3'(f_1)[(g - g_1)df - (f - f_1)dg],$$

that is, is proportional to the Euler form $(g - g_1)df - (f - f_1)dg$. Therefore, the quadratic part of R at (f_1, g_1) is proportional to the exterior product of the differential of the quadratic part of Q at (f_1, g_1) and the Euler form, so it is proportional to the quadratic form of Q at (f_1, g_1) . Thus a suitable linear combination of R and Q has both linear and quadratic parts vanishing at (f_1, g_1) , which implies that the multiplicity of (f_1, g_1) as a solution of the system $Q = R = 0$ is at least 6. □

LEMMA 2.23. *At least six real solutions of $Q = R = 0$ lie outside \mathbb{R}_+^2 .*

Proof. There are exactly four points where the form η_2 vanishes, exactly one in each real quadrant. These points are evidently solutions of the system $Q = R = 0$. Consider the connected component of the curve $\{Q = 0\}$ containing such a point. It is a compact oval not intersecting the coordinate axes. The polynomial R vanishes at least once on this component, namely at this point. Therefore, R should have at least another zero on this oval (counting multiplicities), also necessarily lying in the same quadrant. □

2.3.8. *Final count.* The number N of points in the intersection $\gamma_1 \cap \gamma_2 \cap \mathbb{R}_+^2$ is less than or equal to $2 + 0 + 28 - 18 = 12$, where 2 is the number of unbounded components of γ_1 in \mathbb{R}_+^2 , 0 is the number of unbounded components of $\{Q = 0\}$, 28 is the Kushnirenko–Bernstein upper bound for the number of complex solution of $R = Q = 0$ in $(\mathbb{C}^*)^2$, and 18 is the number of solutions of $R = Q = 0$ outside \mathbb{R}_+^2 counted with multiplicities. Therefore, Proposition 2.9 and Theorem 1.5(b) are finally proved.

2.3.9. *Comments on Theorem 1.5.* (1) Computer experiments indicate that, except for the two solutions of the system $\xi_1 = \xi_2 = \xi_3 = 0$, all the remaining sixteen solutions of the system $R = Q = 0$ can be real. However, not all of them lie in the positive quadrant: typically the system $R = Q = 0$ has four solutions in each real open quadrant. This implies that, provided that this statement about the root configuration could be rigorously proved, the best estimate obtainable by the above method would be 6, very close to Maxwell’s conjectural bound 4.

(2) An even more important observation is that there are typically only two points of intersection of γ_2 and $\Gamma = \{Q = 0\}$ lying in \mathbb{R}_+^2 . In other words, the first application of the Rolle–Khovanskii lemma seems to be exact numerically: a rigorous proof that there are just two points in $\gamma_1 \cap \Gamma \cap \mathbb{R}_+^2$ would imply the original Maxwell conjecture in the case of three charges.

In fact, it is enough to prove a seemingly simpler statement that the number of intersections of Γ and $\gamma_1 = \{f^{1/(\alpha+1)}\xi_2 - \xi_1 = 0\}$ lying in \mathbb{R}_+^2 is at most two. This looks to be easier since the

equation defining γ_1 , being a quadratic polynomial in g , can be solved explicitly. The resulting two solutions $g = g_{1,2}(f)$ parameterize γ_1 , and the problem reduces to the question about the number of positive zeros of a univariate algebraic function $Q(f, g_1(f))$.

(3) The fact that the polynomial R can be reduced to a smaller polynomial \tilde{R} by subtraction of a multiple of Q is a manifestation of a general, yet unexplained, phenomenon: tuples of polynomials resulting from several consecutive applications of the Rolle–Khovanskii theorem are very far from generic, and in every specific case one can usually make a reduction similar to the reduction of R to \tilde{R} above.

2.4. Proof of Theorem 1.7(a)

From now on we will always assume that all our charges are positive (the case of all negative charges follows by a global sign change).

2.4.1. 1-dimensional case. As a warm-up exercise we will prove Theorem 1.7(b) in the simplest case of the x -axis.

The idea of the proof is to use the limit function

$$V_\infty(x) = \min_{i=1,\dots,l} ((x - x_i)^2 + y_i^2) = \lim_{\alpha \rightarrow \infty} V_\alpha^{-1/\alpha}(x), \tag{2.13}$$

where (x_i, y_i) , for $i = 1, \dots, l$, are the coordinates of the i th charge. (We assume for simplicity that all $y_i \neq 0$. The general case follows by taking the limit.) The function $V_\infty(x)$ has at most $l - 1$ points of non-smoothness. Denote these points by γ_j .

LEMMA 2.24. *The convergence $V_\infty(x) = \lim_{\alpha \rightarrow \infty} V_\alpha^{-1/\alpha}(x)$ is valid in the C^2 -class on any closed interval free from the γ_j .*

Proof. We assume that on such an interval $\rho_1 < (1 - \eta)\rho_i$, where $i \geq 2$ and $\eta > 0$. (Here $\rho_i = (x - x_i)^2 + y_i^2$.) Therefore, $V_\infty(x) = \rho_1$ on this interval. The first derivative of $V_\infty(x)$ equals

$$\begin{aligned} (V_\alpha^{-1/\alpha}(x))' &= -\frac{1}{\alpha} V_\alpha^{-1/\alpha-1}(x) \left(-\alpha \sum_{i=1}^l 2\zeta_i \rho_i^{-\alpha-1} (x - x_i) \right) \\ &= 2 \left[\zeta_1(x - x_1) + \sum_{i=2}^l \zeta_i (\rho_i/\rho_1)^{-\alpha-1} (x - x_i) \right] (\rho_1^\alpha V_\alpha(x))^{-2/\alpha-1} \\ &= 2[\zeta_1(x - x_1) + o(1)](\zeta_1 + o(1))^{-2/\alpha-1} \\ &= 2(x - x_1) + o(1) = V'_\infty(x) + o(1), \end{aligned}$$

where $\lim_{\alpha \rightarrow \infty} o(1) = 0$.

Computations with the second derivative are similar, but more cumbersome. □

COROLLARY 2.25. *For sufficiently large α any closed interval free from the γ_j contains at most one critical point of $V_\alpha(x)$.*

Proof. Indeed, for any sufficiently large α the second derivative $(V_\alpha^{-1/\alpha}(x))''$, being close to $V_\infty'' = 2$, is positive on this interval. Therefore, $V_\alpha^{-1/\alpha}(x)$ is convex and can have at most one critical point on this interval. Finally, the critical points of $V_\alpha^{-1/\alpha}(x)$ are the same as the critical points of $V_\alpha(x)$. □

LEMMA 2.26. *For any sufficiently large α a closed interval containing some γ_j and free from the x_i contains at most one critical point of $V_\alpha(x)$.*

Proof. Note that such an interval contains exactly one γ_j since the γ_j are separated by the x_i . The required result follows from the fact that $V_\alpha(x)$ is necessarily convex on any such interval. Indeed,

$$(V_\alpha(x))'' = \alpha(\alpha + 1) \sum_{i=1}^l \zeta_i \rho_i^{-\alpha-2} \left(4(x - x_i)^2 - \frac{2\rho_i}{\alpha + 1} \right).$$

Since $x - x_i \neq 0$ on the interval under consideration, $2\rho_i/(\alpha + 1)$ is necessarily smaller than $4(x - x_i)^2$ for α large enough. Thus, $(V_\alpha(x))''$ is positive (recall that $\zeta_i > 0$) and $V_\alpha(x)$ itself is convex. \square

2.4.2. *Multidimensional case.* We start with the discussion of the critical points of the limiting function $V_\infty(\bar{x})$.

2.4.3. *Critical points of $V_\infty(\bar{x})$.* The function $V_\infty(\bar{x})$ is the only piecewise smooth continuous semialgebraic function. Hence the standard definitions of critical points and Morse indices are not applicable. Therefore we should provide definitions of analogues of critical points of such functions and generalize the notion of Morse index. This is quite a classical subject, and an alternative approach giving the same results can be found in [1, 19].

DEFINITION 2.27. A point \bar{x}_0 is a *critical point* of $V_\infty(\bar{x})$ if for any sufficiently small ball B centered at \bar{x}_0 its subset $B_- = \{V_\infty(\bar{x}) < V_\infty(\bar{x}_0)\} \subseteq B$ is either empty or non-contractible.

The critical point \bar{x}_0 is said to be *non-degenerate* if B_- is either empty or homologically equivalent to a sphere. In this case the *Morse index* of \bar{x}_0 is defined as the dimension of this sphere plus 1. (By default, $\dim(\emptyset) = -1$.)

LEMMA 2.28. *Every effective Voronoi cell of the Voronoi diagram of a generic configuration of positive charges contains a unique critical point of $V_\infty(\bar{x})$. Its index equals the dimension of the Voronoi cell.*

Proof. Indeed, take any effective Voronoi cell S . As above let $\mathcal{NS}(S)$ denote the set of all the nearest sites of S . By definition, S intersects the convex hull of $\mathcal{NS}(S)$. Denote this (unique) intersection point by $p(S)$. We claim that $p(S)$ is the unique critical point of $V_\infty(\bar{x})$ located on S . Indeed, the function $V_\infty(\bar{x})$ restricted to $\mathcal{NS}(S)$ has a local maximum at $p(S)$ since any sufficiently small move within $\mathcal{NS}(S)$ brings us closer to one of the nearest sites. (Here we implicitly use the genericity assumptions on the configuration, that is, that there are exactly $k + 1$ nearest sites for any Voronoi cell of codimension k and that S intersects the interior of the closure of $\mathcal{NS}(S)$.) On the other hand, the restriction of $V_\infty(\bar{x})$ to S itself has the global minimum on S for similar reasons. \square

REMARK. Figure 4, also present in [19], illustrates the above Lemma 2.28. The left picture shows the function $V_\infty(\bar{x}) = \min(\rho_1, \rho_2, \rho_3)$ where the three points are located at $(1, 0)$, $(\pm\frac{1}{2}\sqrt{3}, -\frac{1}{2})$. It is related to the left picture on Figure 1 showing the corresponding potential $V_\alpha(\bar{x})$ for $\alpha = 1$. In this case all the Voronoi cells of the Voronoi diagram are effective and one sees the local maximum inside the convex hull of these points. On the right picture the three points are located at $(0, 0)$, $(2, 0)$ and $(1, \frac{1}{2})$. In this case the 0-dimensional Voronoi cell and one of the 1-dimensional Voronoi cells are ineffective and there is no critical point at the 0-dimensional Voronoi cell. This picture is similarly related to the right picture on Figure 1.

2.4.4. *Proof continued.* In order to settle the multidimensional case we generalize the previous proof using the following idea.

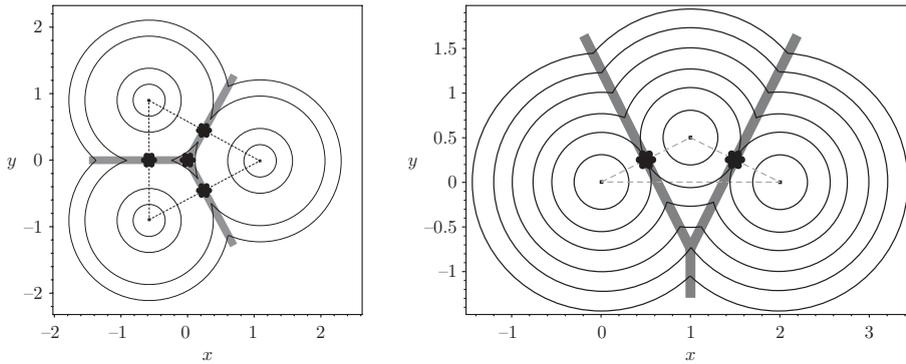


FIGURE 4. *Effective and ineffective 0-dimensional Voronoi cell of $V_\infty(\bar{x})$.*

Main idea for zero-dimensional Voronoi cells. Near an effective zero-dimensional Voronoi cell of the Voronoi diagram the union of the region where $V_\alpha(\bar{x})$ is convex (and therefore has at most one critical point) and the region where $V_\alpha(\bar{x})$ is too C^1 -close to $V_\infty^{-\alpha}(\bar{x})$ to have any critical points asymptotically covers a complete neighborhood of the Voronoi cell. More exactly, we compute asymptotics of the sizes of the above regions, and show that the first region shrinks more slowly than the second region grows.

The following expressions for the gradient and the Hessian form (that is, the quadratic form defined by the matrix of the 2nd partial derivatives) of $V_\alpha(\bar{x})$ are crucial for further computations:

$$\nabla V_\alpha(\bar{x}) = -\alpha \sum_{i=1}^l \zeta_i \rho_i^{-\alpha-1}(\bar{x}) \nabla \rho_i(\bar{x}),$$

$$\text{Hess } V_\alpha(\bar{x}) \cdot \xi = \alpha(\alpha + 1) \sum_{i=1}^l \zeta_i \rho_i^{-\alpha-2}(\bar{x}) \left((\nabla \rho_i(\bar{x}), \xi)^2 - \frac{2}{\alpha + 1} \rho_i(\bar{x}) \|\xi\|^2 \right).$$

Here \cdot denotes the evaluation of the quadratic form $\text{Hess } V_\alpha(\bar{x})$ at ξ .

We start with the case of zero-dimensional Voronoi cells. The general case will be treated as the direct product of the zero-dimensional case in the direction transversal to the Voronoi cell and the full-dimensional case along the Voronoi cell.

2.4.5. Zero-dimensional Voronoi cells of a Voronoi diagram. Let S be a zero-dimensional Voronoi cell of a Voronoi diagram of a generic configuration. We can assume that

$$\rho_1(S) = \dots = \rho_{n+1}(S) < \rho_i(S) \quad \text{for } i > n + 1.$$

Define the functions

$$m(\bar{x}) = \min_{i=1, \dots, n+1} \rho_i(\bar{x}), \quad M(\bar{x}) = \max_{i=1, \dots, n+1} \rho_i(\bar{x}), \quad r(\bar{x}) = \min_{i=n+2, \dots, l} \rho_i(\bar{x}),$$

and let

$$\phi(\bar{x}) = \log \left(\frac{M(\bar{x})}{m(\bar{x})} \right).$$

Note that $\phi(\bar{x})$ is everywhere positive except at the origin, and is equivalent to the Euclidean distance to S in a sufficiently small neighborhood of S .

LEMMA 2.29. *There exists a $\delta > 0$ such that for any \bar{x} in the δ -neighborhood \mathcal{U} of S the following conditions hold.*

(i) There exists an $\epsilon > 0$ such that, for any $\bar{x} \in \mathcal{U}$,

$$r(\bar{x}) > e^{2\epsilon} m(\bar{x}) > e^\epsilon M(\bar{x})$$

(in particular, $\phi(\bar{x}) < \epsilon$ in \mathcal{U}).

(ii) Absolute values of all ratios $c_k(\bar{x})/c_l(\bar{x})$ of the unique linear relation

$$\sum_{i=1}^{n+1} c_i(\bar{x}) \nabla \rho_i(\bar{x}) = 0$$

are bounded by some constant $\Upsilon > 0$.

(iii) If the cell S is ineffective, then the closure of \mathcal{U} can be separated from the convex hull of $\mathcal{NS}(S)$ by a hyperplane.

Proof. For the point S itself, $r(S) > m(S) = M(S)$. Also, in the unique relation

$$\sum_{i=1}^{n+1} c_i(\bar{x}) \nabla \rho_i(S) = 0$$

all coefficients are non-zero by the genericity of the configuration. Therefore both claims follow from continuity of the involved functions. \square

We prove that for a sufficiently large α the domain \mathcal{U} is the union of two subdomains $\mathcal{U} = U_1 \cup U_2$ such that $V_\alpha(\bar{x})$ is convex in U_1 and has no critical points in U_2 .

In what follows C_k and κ_k denote positive constants independent of α but dependent on the configuration and the choice of \mathcal{U} .

LEMMA 2.30. *There exists a constant κ_1 independent of α such that, for α sufficiently large, the function $V_\alpha(\bar{x})$ has no critical points in the domain defined by $\{\phi(\bar{x}) > \kappa_1/(\alpha + 1)\} \cap \mathcal{U}$.*

Proof. First, consider the case $l = n + 1$.

The condition $\nabla V_\alpha(\bar{x}) = 0$ implies that

$$\frac{\zeta_i}{\zeta_j} \left(\frac{\rho_i}{\rho_j} \right)^{-\alpha-1} = \frac{c_j}{c_i} \leq \Upsilon,$$

and, taking the logarithm of both sides, we arrive at

$$\phi(\bar{x}) < \frac{\log \Upsilon + \log \max_{i,j} (c_i/c_j)}{\alpha + 1}.$$

Thus one can take $\kappa_1 = \log \Upsilon + \log \max_{i,j} (c_i/c_j)$ in this case.

The case $l > n + 1$ differs by exponentially small terms. Namely, suppose that $\rho_1(\bar{x}) = M(\bar{x})$. Then

$$0 = \rho_1^{\alpha+1} \nabla V_\alpha(\bar{x}) = (\zeta_1 \nabla \rho_1 - \xi) + \sum_{i=2}^{n+1} \zeta_i \left(\frac{\rho_i}{\rho_1} \right)^{-\alpha-1} \nabla \rho_i,$$

where $\xi = \sum_{i=n+2}^l \zeta_i (\rho_i/\rho_1)^{-\alpha-1} \nabla \rho_i$. One can easily see that $\|\xi\| \leq C_1 e^{-\epsilon(\alpha+1)}$. Therefore, since $\nabla \rho_2(x), \dots, \nabla \rho_{n+1}(x)$ are linearly independent in \mathcal{U} ,

$$|\zeta_i (\rho_i/\rho_1)^{-\alpha-1} - \zeta_1 c_i/c_1| \leq C_2 e^{-\epsilon(\alpha+1)} = o(1),$$

and we get the required estimate. \square

LEMMA 2.31. *There exists a constant κ_2 independent of α such that for all sufficiently large α the function $V_\alpha(\bar{x})$ is convex in the domain $\{\phi(\bar{x}) < (\log(\alpha + 1) - \kappa_2)/(\alpha + 2)\}$.*

Proof. Again, start with the case $l = n + 1$. The gradients $\nabla\rho_i$, for $i = 1, \dots, n + 1$, span the whole \mathbb{R}^n . Thus, the quadratic form $\sum_{i=1}^{n+1} \zeta_i (\nabla\rho_i \cdot \xi)^2 \geq C_3 \|\xi\|^2 > 0$ is positive definite. Therefore, one gets

$$\begin{aligned} \frac{1}{\alpha(\alpha + 1)} \text{Hess } V_\alpha(\bar{x}) \cdot \xi &= \sum_{i=1}^{n+1} \zeta_i \rho_i^{-\alpha-2} (\nabla\rho_i, \xi)^2 - \frac{2}{\alpha + 1} \sum_{i=1}^{n+1} \zeta_i \rho_i^{-\alpha-1} \|\xi\|^2 \\ &\geq M(\bar{x})^{-\alpha-2} \left(\sum_{i=1}^{n+1} \zeta_i (\nabla\rho_i, \xi)^2 \right) - \frac{2(n + 1) \max \zeta_i}{\alpha + 1} m(\bar{x})^{-\alpha-1} \|\xi\|^2 \\ &\geq m(\bar{x})^{-\alpha-2} \left(C_3 \cdot (e^{\phi(\bar{x})})^{-\alpha-2} - \frac{C_4}{\alpha + 1} \right) \|\xi\|^2. \end{aligned}$$

The last form is positive definite if

$$e^{-(\alpha+2)\phi(\bar{x})} > \frac{C_5}{\alpha + 1} \quad \text{or, equivalently,} \quad \phi(\bar{x}) < \frac{\log(\alpha + 1) - \kappa_2}{\alpha + 2}. \tag{2.14}$$

The case $l > n + 1$ differs by an exponentially small term, namely by the term

$$\left| \sum_{i=n+2}^l \zeta_i \rho_i^{-\alpha-2} \left[(\nabla\rho_i, \xi)^2 - \frac{2\rho_i}{\alpha + 1} \|\xi\|^2 \right] \right| \leq C_6 m(\bar{x})^{-\alpha-2} e^{-\epsilon(\alpha+2)} \|\xi\|^2.$$

Therefore, instead of (2.14) we find that $\text{Hess } V$ is positive definite provided

$$e^{-(\alpha+2)\phi(\bar{x})} > \frac{C_5}{\alpha + 1} + C_6 e^{-\epsilon(\alpha+2)},$$

which gives the same estimate with a different constant. □

LEMMA 2.32. *The function $V_\alpha(\bar{x})$ has at most one critical point in \mathcal{U} . If the cell under consideration is effective then the critical point exists and is a local minimum. If the cell under consideration is not effective then there is no critical point in \mathcal{U} .*

Proof. Indeed, in the above notation, for sufficiently large α one has

$$\frac{\log(\alpha + 1) - \kappa_2}{\alpha + 2} > \frac{\kappa_1}{\alpha + 1}.$$

Thus \mathcal{U} is covered by two domains,

$$\left\{ \phi(\bar{x}) > \frac{\kappa_1}{\alpha + 1} \right\} \quad \text{and} \quad \left\{ \phi(\bar{x}) \leq \frac{\log(\alpha + 1) - \kappa_2}{\alpha + 2} \right\}.$$

By Lemma 2.30, $V_\alpha(\bar{x})$ has no critical points in the first domain. By Lemma 2.31, $V_\alpha(\bar{x})$ is convex and has at most one critical point in the second domain.

In the case when the considered 0-dimensional Voronoi cell is effective, $V_\alpha(\bar{x})$ actually has a local minimum located close to that Voronoi cell: the function $V_\alpha^{-1/\alpha}(\bar{x})$, being C^0 -close to V_∞ , has a local minimum inside \mathcal{U} .

The last statement is a particular case of Lemma 2.36 below. □

Taken together with Lemmas 2.30 and 2.31, this proves that, for α sufficiently large, to each effective zero-dimensional Voronoi cell of a generic configuration of points one has exactly one corresponding minimum of V_α .

2.4.6. Case of arbitrary codimension. Let S be any Voronoi cell of codimension k of the Voronoi diagram. We prove that for a generic configuration of positive charges and any compact $K \subset S$ lying inside S there exists a sufficiently small neighborhood \mathcal{U}_K independent of α and

containing at most one critical point of $V_\alpha(\bar{x})$. Moreover, this critical point exists if and only if the cell is effective, and its Morse index is equal to $n - k$.

Denote by L the affine subspace spanned by S . Recall that the first genericity assumption means that there exist exactly $k + 1$ charges $\zeta_1, \dots, \zeta_{k+1}$ closest to S . Denote by M the affine subspace orthogonal to L .

LEMMA 2.33. $\dim M = k$.

Proof. Indeed, a small shift of any point of S in any direction orthogonal to M produces a point with the same set of closest charges: distances to the charges not in $\mathcal{NS}(S)$ will still remain bigger than the distances to the charges in $\mathcal{NS}(S)$, and the latter distances will remain equal. Therefore the shifted point still lies in S , so the dimension of S is at least $\text{codim } M$, that is, $\dim M \geq k$. The opposite inequality is evident since $\mathcal{NS}(S)$ contains $k + 1$ points. \square

If the Voronoi cell S intersects the convex hull of $\mathcal{NS}(S)$ then the second genericity assumption means that any k of the charges in $\mathcal{NS}(S)$ do not lie on a hyperplane in M passing through the point $L \cap M$.

Choosing an appropriate coordinate system we may assume that L and M intersect at the origin, that is, are orthogonal complements of each other. Denote by \bar{x}_L and \bar{x}_M orthogonal projections of a vector \bar{x} to the linear subspaces L and M respectively; thus $\bar{x} = \bar{x}_M + \bar{x}_L$.

We assume that the set $\mathcal{NS}(S)$ of charges closest to S consists of charges $\zeta_1, \dots, \zeta_{k+1}$ and denote the distances to these charges by $\rho_1, \dots, \rho_{k+1}$ respectively.

As before, introduce the functions

$$m(\bar{x}) = \min_{i=1, \dots, k+1} \rho_i(\bar{x}), \quad M(\bar{x}) = \max_{i=1, \dots, k+1} \rho_i(\bar{x}), \quad r(\bar{x}) = \min_{i=k+2, \dots, l} \rho_i(\bar{x}),$$

and let

$$\phi_M(\bar{x}) = \log \left(\frac{M(\bar{x})}{m(\bar{x})} \right).$$

Let K be a compact subset of S .

LEMMA 2.34. *There exists $\delta > 0$ so small that in the δ -neighborhood $\mathcal{U}_K \subset \mathbb{R}^n$ of K the following conditions hold.*

(i) *There exists an ϵ with $0 < \epsilon$ such that for any $\bar{x} \in \mathcal{U}_K$ one has*

$$r(\bar{x}) > e^{2\epsilon} m(\bar{x}) > e^\epsilon M(\bar{x}).$$

(ii) *The absolute value of all ratios $c_k(\bar{x})/c_l(\bar{x})$ in the unique linear dependence*

$$\sum_{j=1}^{k+1} c_k(\bar{x}) \nabla_M \rho_j = 0$$

is bounded from above by some constant Υ , where $\nabla_M \rho_j$ denotes the orthogonal projection of the gradient $\nabla \rho_j$ to M . Note that the tuple of $c_k(\bar{x})$ is, up to proportionality, constant on S (and none of the c_k vanishes due to the genericity assumptions).

(iii) *If the cell S is ineffective, then the closure of \mathcal{U} can be separated from the convex hull of $\mathcal{NS}(S)$ by a hyperplane.*

Proof. The proof is analogous to the proof of Lemma 2.29: the required relations hold on S itself, and can be extended to a neighborhood of K by continuity. \square

Note that ϕ_M is equivalent to $\|\bar{x}_M\|$ near the origin:

$$C_M^{-1} \phi_M(\bar{x}) \leq \|\bar{x}_M\| \leq C_M \phi_M(\bar{x}) \tag{2.15}$$

for some $C_M > 0$ and all $x \in \mathcal{U}_K$.

We claim that for α big enough the critical points of $V_\alpha(\bar{x})$ in \mathcal{U}_K can occur only in a small neighborhood of the origin, shrinking as $\alpha \rightarrow \infty$.

LEMMA 2.35. (a) For a certain positive κ_3 the function $V_\alpha(\bar{x})$ has no critical points in the domain given by $\mathcal{U}_K \cap \{\|\bar{x}_L\| > \kappa_3 \cdot e^{-\epsilon(\alpha+1)}\}$.

(b) For a certain positive κ_4 the function $V_\alpha(\bar{x})$ has no critical points in the domain given by $\mathcal{U}_K \cap \{\phi_M(\bar{x}) > \kappa_4/(\alpha + 1)\}$.

Proof. (a) The idea is that outside M the gradients of the ρ_i , for $i = 1, \dots, k + 1$, are all directed away from M . The contribution of the remaining ρ_i , being exponentially small, is negligible outside an exponentially small neighborhood.

We calculate the directional derivative of $V_\alpha(\bar{x})$ in the direction \bar{x}_L at a point $\bar{x} \in \mathcal{U}_K$:

$$-\frac{1}{\alpha} \frac{\partial V_\alpha(\bar{x})}{\partial \bar{x}_L}(\bar{x}) = \sum_{i=1}^{k+1} \zeta_i \rho_i^{-\alpha-1} \frac{\partial \rho_i}{\partial \bar{x}_L}(\bar{x}) + \sum_{i=k+2}^l \zeta_i \rho_i^{-\alpha-1} \frac{\partial \rho_i}{\partial \bar{x}_L}(\bar{x}).$$

Since

$$\frac{\partial \rho_i}{\partial \bar{x}_L}(\bar{x}) = 2\|\bar{x}_L\|^2 \quad \text{for } i = 1, \dots, k + 1,$$

we conclude that the absolute value of the first term is at least $C_9 M(\bar{x})^{-\alpha-1} \|\bar{x}_L\|^2$. The absolute value of the second term is at most

$$C_{10} r(\bar{x})^{-\alpha-1} \|\bar{x}_L\| < C_{10} M(\bar{x})^{-\alpha-1} e^{-\epsilon(\alpha+1)} \|\bar{x}_L\|,$$

and the estimate follows.

(b) We essentially repeat the computations of Lemma 2.30:

$$\begin{aligned} -\frac{1}{\alpha} \nabla_M V_\alpha(\bar{x}) &= \sum_{i=1}^{k+1} \zeta_i \rho_i^{-\alpha-1} \nabla_M \rho_i(\bar{x}) + \sum_{j=k+2}^l \zeta_j \rho_j^{-\alpha-1} \nabla_M \rho_j(\bar{x}) \\ &= M(\bar{x})^{-\alpha-1} \left[\sum_{i=1}^{k+1} \zeta_i \left(\frac{\rho_i}{M(\bar{x})} \right)^{-\alpha-1} \nabla_M \rho_i(\bar{x}) + O(e^{-\epsilon(\alpha+1)}) \right]. \end{aligned}$$

So $\nabla_M V_\alpha(\bar{x}) = 0$ implies, as in Lemma 2.30, that the functions $(\rho_i/M(\bar{x}))^{-\alpha-1}$ are bounded, which gives $\phi_M(\bar{x}) < \kappa_4/(\alpha + 1)$. □

COROLLARY 2.36. If a cell S is not effective then for α large enough the function $V_\alpha(\bar{x})$ has no critical points in \mathcal{U}_K .

Proof. Indeed, in this case $\mathcal{U}_K \subset \{\|\bar{x}_L\| > c\}$ for some $c > 0$. □

From now on we suppose that K and the convex hull of $\mathcal{NS}(S)$ intersect at the origin, and consider the domain

$$\mathcal{U}_\alpha = \left\{ \bar{x}_L < \kappa_3 \cdot e^{-\epsilon(\alpha+1)}, \|\bar{x}_M\| < \frac{C_M \kappa_4}{\alpha + 1} \right\} \subset \mathcal{U}_K. \tag{2.16}$$

The union of this domain and the domain described in Lemma 2.35 covers \mathcal{U} by (2.15). Moreover, $\phi_M(\bar{x}) \leq C_M^2 \kappa_4/(\alpha + 1)$ in \mathcal{U} .

Our next goal is to study the quadratic form $\text{Hess } V_\alpha(\bar{x})$ in the domain \mathcal{U}_α .

LEMMA 2.37. Let $V_{k,\alpha}(\bar{x}) = \sum_{i=1}^{k+1} \zeta_i \rho_i^{-\alpha}$. The following hold.

(a) We have $\text{Hess } V_\alpha(\bar{x}) - \text{Hess } V_{k,\alpha}(\bar{x}) = M(\bar{x})^{-\alpha-2} O(e^{-\epsilon\alpha})$ for $\bar{x} \in \mathcal{U}_\alpha$ and $\alpha \rightarrow \infty$.

(b) *There exist two quadratic forms $A(x)$ and $B(x)$ on M and L respectively such that $\text{Hess } V_{k,\alpha}(\bar{x}) = A(\bar{x}) \oplus B(\bar{x}) + M(\bar{x})^{-\alpha-2}O(e^{-\epsilon\alpha})$.*

In addition, there exist some positive constants κ_5 and κ_6 such that the form A is positive definite and bounded from below by a $\kappa_5 \cdot \alpha^2 M(\bar{x})^{-\alpha-2}$, and the form B is negative definite and bounded from above by $-\kappa_6 \cdot \alpha M(\bar{x})^{-\alpha-2}$.

Proof. (a) We have to estimate from above the contribution of the distant charges:

$$\begin{aligned} |\text{Hess } V_\alpha(\bar{x}) \cdot \xi - \text{Hess } V_{k,\alpha}(\bar{x}) \cdot \xi| &\leq \alpha(\alpha + 1) \sum_{i=k+2}^l \zeta_i \rho_i^{-\alpha-2} \left((\nabla \rho_i, \xi)^2 - \frac{2}{\alpha + 1} \rho_i \|\xi\|^2 \right) \\ &\leq CM(\bar{x})^{-\alpha-2} e^{-\epsilon(\alpha+2)} \|\xi\|^2. \end{aligned}$$

(b) For $i = 1, \dots, k + 1$ the charges ζ_i are in M . Therefore we have

$$\begin{aligned} (\nabla \rho_i(\bar{x}), \xi)^2 &= (\nabla_M \rho_i(\bar{x}), \xi_M)^2 + 2(\nabla_M \rho_i(\bar{x}), \xi_M)(\nabla_L \rho_i(\bar{x}), \xi_L) + (\nabla_L \rho_i(\bar{x}), \xi_L)^2 \\ &= (\nabla_M \rho_i(\bar{x}), \xi_M)^2 + O(e^{-\epsilon\alpha}) \|\xi\|^2. \end{aligned}$$

Here we used the facts that $\nabla_L \rho_j(\bar{x}) = 2\bar{x}_L$ and $\|\bar{x}_L\| \leq \kappa_3 e^{-\epsilon\alpha}$ in \mathcal{U}_α .

Therefore,

$$\begin{aligned} \text{Hess } V_{k,\alpha}(\bar{x}) \cdot \xi &= \alpha(\alpha + 1) \sum_{i=1}^{k+1} \zeta_i \rho_i^{-\alpha-2} \left[(\nabla \rho_i, \xi)^2 - \frac{2}{\alpha + 1} \rho_i \|\xi\|^2 \right] \\ &= \alpha(\alpha + 1) \sum_{i=1}^{k+1} \zeta_i \rho_i^{-\alpha-2} (\nabla_M \rho_i, \xi_M)^2 - 2\alpha \sum_{i=1}^{k+1} \zeta_i \rho_i^{-\alpha-1} \|\xi\|^2 \\ &\quad + O(e^{-\epsilon\alpha}) \sum_{i=1}^{k+1} \rho_i^{-\alpha-2} \|\xi\|^2. \end{aligned}$$

Due to the linear independence of the $\nabla_M \rho_j$ in \mathcal{U}_α we get, exactly as in Lemma 2.31,

$$\sum_{i=1}^l \zeta_i (\nabla_M \rho_i, \xi_M)^2 > C_9 \|\xi_M\|^2$$

for some $C_9 > 0$. Therefore, one can estimate the first term from below as

$$\alpha(\alpha + 1) \sum_{i=1}^{k+1} \zeta_i \rho_i^{-\alpha-2} (\nabla \rho_j, \xi_M)^2 \geq \alpha(\alpha + 1) C_9 M(\bar{x})^{-\alpha-2} \|\xi_M\|^2.$$

The second term can be estimated using

$$(\min \zeta_i) M(\bar{x})^{-\alpha-1} \leq \frac{1}{k + 1} \sum_{j=1, \dots, k+1} \zeta_i \rho_j^{-\alpha-1} \leq (\max \zeta_i) M(\bar{x})^{-\alpha-1} e^{(\alpha+1)\phi_M(\bar{x})}.$$

The expression $e^{(\alpha+1)\phi_M(\bar{x})}$ is bounded in \mathcal{U}_α by a constant $C_{10} = e^{C_M^2 \kappa_4}$. Therefore the restriction of $\text{Hess } V_{k,\alpha}(\bar{x}) \cdot \xi$ to M has the lower bound

$$\alpha(\alpha + 1) M(\bar{x})^{-\alpha-2} \left(C_9 - \frac{2(k + 1)C_{10}}{\alpha + 1} \max_{i, \bar{x} \in \mathcal{U}} \zeta_i \rho_i(\bar{x}) - O(e^{-\epsilon\alpha}) \right) \|\xi_M\|^2.$$

On the other hand, the restriction of $\text{Hess } V_{k,\alpha}(\bar{x}) \cdot \xi$ to L is negative definite and has the upper bound

$$-2\alpha(k + 1) \min \zeta_i M(\bar{x})^{-\alpha-1} \|\xi_L\|^2 (1 - O(e^{-\epsilon\alpha})).$$

COROLLARY 2.38. *For large enough α and any $\bar{x} \in \mathcal{U}_\alpha$ the signature of the quadratic form $\text{Hess } V_\alpha(\bar{x})$ is $(k, n - k)$.*

LEMMA 2.39. *For sufficiently large α there is at most one critical point of $V_\alpha(\bar{x})$ in \mathcal{U}_K , and its Morse index is equal to $n - k$.*

Proof. As was shown in Lemma 2.35, there are no critical points of V_α in $\mathcal{U}_K \setminus \mathcal{U}_\alpha$ for α large enough. So it is enough to prove that the mapping dV_α is one-to-one in \mathcal{U}_α .

Take any segment $I = \{a_t = a + t\xi : 0 \leq t \leq 1\} \subset \mathcal{U}_\alpha$. Let $\pi(t)$ be the projection of the point $dV_\alpha(a_t)$ on the direction $\bar{\xi} = \xi_M - \xi_L$ (where $\xi = \xi_M + \xi_L$). We claim that $\pi(t)$ is a monotone function of t , and, therefore, its values at a_0 and a_1 cannot coincide. Indeed, using Lemma 2.37 one can estimate $\pi'(t)$ from below as

$$\begin{aligned} \pi'(t) &= \left(\bar{\xi}, \frac{\partial}{\partial t} dV_\alpha(a_t) \right) = \text{Hess } V_\alpha(\bar{\xi}, \xi) \\ &\geq M(\bar{x})^{-\alpha-2} [\kappa_5 \cdot \alpha^2 \|\xi_M\|^2 + \kappa_6 \cdot \alpha \|\xi_L\|^2 + O(e^{-\epsilon\alpha}) \|\xi\|^2] > 0 \end{aligned}$$

for α sufficiently large.

This, by convexity of \mathcal{U}_α , immediately implies the claim of the lemma: assuming that there are two critical points and joining them by a segment we get a contradiction. \square

In Lemma 2.36 we proved that ineffective cells do not create critical points of V_α for α large enough. To finish the proof of the one-to-one correspondence between effective cells and critical points of V_α we have to show that if K contains a critical point of the function V_∞ then \mathcal{U}_K does contain a critical point of V_α .

LEMMA 2.40. *Assume that K contains a critical point of V_∞ . Then \mathcal{U}_K contains a critical point of V_α for α sufficiently large.*

The idea of the proof is that the smooth function $V_\alpha^{-1/\alpha}$ is arbitrarily \mathcal{C}^0 -close to V_∞ for large α . The topology of the sets $X_c = \{V_\infty \leq c\} \cap \mathcal{U}_K$ changes as c passes the critical value implying the change of the topology of the sets $Y_c = \{V_\alpha^{-1/\alpha} \leq c\} \cap \mathcal{U}_K$ for sufficiently large α , which in its turn implies the appearance of the critical points of $V_\alpha^{-1/\alpha}$, the latter coinciding with the critical points of V_α .

Let c_0 be the critical value of V_∞ at the point $c \in K$, and let c_1 and c_2 be some regular values, $\min_{\mathcal{U}_K} V_\infty < c_1 < c_0 < c_2 < \max_{\mathcal{U}_K} V_\infty$. We assume that c_1 and c_2 are so close to c_0 that the interval $[c_1, c_2]$ contains no critical values of V_∞ restricted to the boundary of \mathcal{U}_K .

Let $\delta \ll (c_0 - c_1)/2$. For α large enough the function $V_\alpha^{-1/\alpha}$ is at least $\delta/2$ -close to V_∞ . Thus, we have

$$X_{c_1} \subset Y_{c_1+\delta/2} \subset X_{c_1+\delta} \subset Y_{c_1+3\delta/2}.$$

These inclusions induce homomorphisms in homology groups:

$$H_*(X_{c_1}) \longrightarrow H_*(Y_{c_1+\delta/2}) \longrightarrow H_*(X_{c_1+\delta}) \longrightarrow H_*(Y_{c_1+3\delta/2}). \tag{2.17}$$

The composition of the first two homomorphisms is a homomorphism induced by the inclusion $X_{c_1} \subset X_{c_1+\delta}$, which is an isomorphism. Therefore the middle homomorphism in (2.17) is surjective.

Similarly, the composition of the last two homomorphisms is a homomorphism induced by the inclusion $Y_{c_1+\delta/2} \subset Y_{c_1+3\delta/2}$, which, assuming that $V_\alpha^{-1/\alpha}$ has no critical points, is an isomorphism as well. Therefore the middle homomorphism in (2.17) is injective.

Summing up, we conclude that the middle homomorphism is an isomorphism, and Y_{c_1} is homologically equivalent to X_{c_1} .

Similarly, Y_{c_2} is homologically equivalent to X_{c_2} . However, the sets X_{c_1} and X_{c_2} are not homologically equivalent: the first is homologically equivalent to a sphere, and the second is

contractible. Therefore Y_{c_1} and Y_{c_2} are also homologically different implying that $V_\alpha^{-1/\alpha}$ should have a critical point in \mathcal{U}_K .

2.4.7. *Completing the proof of Theorem 1.7(a).* Since V_α has no critical points outside the convex hull of the charges, it is enough to consider, instead of \mathbb{R}^n , an open ball B containing all charges.

It is easy to see that one can cover B by neighborhoods \mathcal{U}_k as in Lemmas 2.29 and 2.34. Indeed, start from the zero-dimensional Voronoi cells S_i , and choose their neighborhoods \mathcal{U}_i according to Lemma 2.29. Then choose compact K_i^1 in one-dimensional Voronoi cells in such a way that their union with these neighborhoods covers the intersection of the union of all one-dimensional Voronoi cells with B . Choose neighborhoods $\mathcal{U}_{K_i^1}$ of these compact subsets according to Lemma 2.34. Then choose compact subsets K_i^2 of two-dimensional Voronoi cells in such a way that $(\bigcup K_i^2) \cup (\bigcup \mathcal{U}_{K_i^1})$ covers the intersection of the union of all two-dimensional Voronoi cells with B , and so on. At the end the union of all selected neighborhoods will cover B .

Each of the neighborhoods corresponding to effective Voronoi cells will contain one critical point of V_α , and its Morse index will be equal to the dimension of the Voronoi cell. The neighborhoods of ineffective Voronoi cells will not contain critical points of V_α .

2.5. *Proof of Theorem 1.7(b)*

We are looking for the critical points of the function \tilde{V}_α defined in a linear space L ,

$$\tilde{V}_\alpha = \sum \zeta_i \tilde{\rho}_i^{-\alpha}, \quad \text{where } \tilde{\rho}_i = \text{dist}^2(\bar{x}, \tilde{c}_i) + y_i^2,$$

where \tilde{c}_i are now *orthogonal projections* on L of the positions c_i of the charges ζ_j . The claim is that the critical points of \tilde{V}_α are in one-to-one correspondence with the cells of the Voronoi diagram of $\{\zeta_j\}$ that are effective with respect to L .

We can characterize the partition of L by intersections with the cells of the Voronoi diagram only in terms of $\tilde{\rho}_i$. Namely, a generalized Voronoi diagram in L is defined as the classical Voronoi diagram in §1.1, with ρ_i replaced by $\tilde{\rho}_i$: a cell S of a generalized Voronoi diagram is the set of all points $\bar{x} \in N$ with the same set $\mathcal{NS}(S) = \{i : \text{for all } k, \tilde{\rho}_i(\bar{x}) \leq \tilde{\rho}_k(\bar{x})\}$. One can immediately see that, thus defined, the generalized Voronoi diagram coincides with the intersection of the original Voronoi diagram with L .

It turns out that using this notation one can get the proof of Theorem 1.7(b) from that of Theorem 1.7(a) by a simple replacement of \mathbb{R}^n by L , ρ_j by $\tilde{\rho}_j$, the charges ζ_j by their projections on L , ∇V_α by $\nabla \tilde{V}_\alpha$, and cells of the Voronoi diagram by cells of the generalized Voronoi diagram. Namely, since $\tilde{\rho}_j$ is just the square of the distance to \tilde{c}_j (up to a constant), exactly the same formulae for $\nabla \tilde{V}_\alpha$ and $\text{Hess } \tilde{V}_\alpha$ hold. Since the $\tilde{\rho}_j$ are radially symmetric, the condition that L intersects the Voronoi diagram generically implies the linear independence of $\nabla \tilde{\rho}_j$, which guarantees the second property of \mathcal{U} and \mathcal{U}_K , etc.

The only difference appears for the full-dimensional strata, that is, the case of $k = 0$ in Lemmas 2.35 and 2.37: while in Theorem 1.7(a) these strata do not contain critical points, in Theorem 1.7(b) the full-dimensional strata corresponding to strictly positive $\tilde{\rho}_j$ will have a critical point. This point will be necessarily unique and is a local maximum by Lemmas 2.35 and 2.37 (modified as mentioned above).

We leave it as an exercise to check that the aforementioned modifications of the proof of Theorem 1.7(a) produce a correct proof of Theorem 1.7(b).

2.6. *Proof of Theorem 1.10*

Denote by a_α^j the number of critical points of Morse index j of $V_\alpha(\bar{x})$. The standard potential $V_1(\bar{x})$ is harmonic in \mathbb{R}^3 , and, therefore, has no local maxima or minima, that is, $a_1^0 = 0$. Using

the Euler characteristics one can easily check that $a_\alpha^2 - l + 1 = a_\alpha^1 - a_\alpha^0$ for any α . Therefore, for $\alpha = 1$ the total number of critical points $a_1^2 + a_1^1 + a_1^0$ equals $2a_1^2 - l + 1$. By the Maxwell inequalities (1.4) one gets $a_1^2 \leq a_\infty^2 \leq \frac{1}{2}l(l-1)$; see §1. Thus,

$$a_1^2 + a_1^1 + a_1^0 = 2a_1^2 - l + 1 \leq l(l-1) - l + 1 = (l-1)^2,$$

exactly Maxwell's estimate.

3. Remarks and problems

REMARK 1. The main objects of consideration in this paper have a strong resemblance with the main objects in tropical algebraic geometry. Namely, the potential $V_\alpha(\bar{x})$ resembles an actual algebraic hypersurface while $V_\infty(\bar{x})$ resembles its tropical limit. Also, Voronoi diagrams are piecewise linear objects as well as tropical curves. Is this a pure coincidence?

REMARK 2. What happens in the case of charges of different magnitudes and signs? Note that in a Voronoi cell of highest dimension corresponding to a negative charge the potential of this charge outweighs potentials of all other charges for large α , and $|V_\alpha|^{-1/\alpha}$ converges uniformly on compact subsets of this cell to $V_\infty(\bar{x})$. Therefore it seems that the function defined on the union of Voronoi cells of highest dimension as

$$\tilde{V}_\infty(\bar{x}) = \text{sign } \zeta_i \cdot \rho_i(x) \quad \text{if } \rho_i(x) = \min_j \rho_j(x)$$

is responsible for the critical points of V_α as $\alpha \rightarrow \infty$.

REMARK 3. Theorem 1.7 is similar to the results of Varchenko [20] and Orlik and Terao [16] on the number of critical points for the product of powers of real linear forms and the number of open components in the complement to the corresponding arrangement of affine hyperplanes. Is there an appropriate result?

REMARK 4. Conjecturally in the case of unit charges the number of critical points of $V_\alpha(\bar{x})$ is bounded from above by the number of effective Voronoi cells in the corresponding Voronoi diagram. The number of all Voronoi cells in Voronoi diagrams in \mathbb{R}^n with l sites has a nice upper bound. What is the upper bound for the number of effective Voronoi cells? Is it the same as for all Voronoi cells?

REMARK 5. Many statements in the paper are valid if one substitutes the potential $r^{-\alpha}$ of a unit charge located at the origin by more or less any concave function $\psi(r)$ of the radius in \mathbb{R}^n . To what extent can the above results and conjecture be generalized for $\psi(r)$ -potentials?

REMARK 6. The initial hope in settling Conjecture 1.8 was related to the fact that in our numerical experiments for a fixed configuration of charges the number of critical points of $V_\alpha(\bar{x})$ was a *non-decreasing* function of α . Unfortunately this monotonicity turned out to be wrong in the most general formulation: the number of critical points of a restriction of a potential to a line is not a monotonic function of α .

EXAMPLE 3.1. The potential

$$V_\alpha(x) = [(x+30)^2 + 25]^{-\alpha} + [(x+20)^2 + 49]^{-\alpha} + [(x+2)^2 + 144]^{-\alpha} \\ + [(x-20)^2 + 49]^{-\alpha} + [(x-30)^2 + 25]^{-\alpha}$$

has three critical points for $\alpha = 0.1$, seven critical points for $\alpha = 0.2$, again three critical points for $\alpha = 0.3$, and again seven critical points as $\alpha = 1.64$, and nine critical points for $\alpha \geq 1.7$.

The existence of such an example for the potential itself (and not of its restriction) is unknown.

Appendix. *James C. Maxwell on points of equilibrium*

In his monumental Treatise [12] Maxwell foresaw the development of several mathematical disciplines. In the passage which we have the pleasure to present to the readers his arguments are that of Morse theory developed at least 50 years later. He uses the notion of periphractic number, or degree of periphraxy which is the rank of H_2 of a domain in \mathbb{R}^3 defined as the number of interior surfaces bounding the domain, and the notion of cyclomatic number, or degree of cyclosis which is the rank H_1 of a domain in \mathbb{R}^3 defined as the number of cycles in a curve obtained by a homotopy retraction of the domain (none of these notions rigorously existed then). (For definitions of these notions see [12, Section 18].) Then he actually proves Theorem 1.1 of §1 which is usually attributed to M. Morse. Finally in Section 113 he makes the following claim.

To determine the number of the points and lines of equilibrium, let us consider the surface or surfaces for which the potential is equal to C , a given quantity. Let us call the regions in which the potential is less than C the negative regions, and those in which it is greater than C the positive regions. Let V_0 be the lowest and V_1 the highest potential existing in the electric field. If we make $C = V_0$ the negative region will include only one point or conductor of the lowest potential, and this is necessarily charged negatively. The positive region consists of the rest of the space, and since it surrounds the negative region it is periphractic.

If we now increase the value of C , the negative region will expand, and new negative regions will be formed round negatively charged bodies. For every negative region thus formed the surrounding positive region acquires one degree of periphraxy.

As the different negative regions expand, two or more of them may meet at a point or a line. If $n + 1$ negative regions meet, the positive region loses n degrees of periphraxy, and the point or the line in which they meet is a point or line of equilibrium of the n th degree.

When C becomes equal to V_1 the positive region is reduced to the point or the conductor of highest potential, and has therefore lost all its periphraxy. Hence, if each point or line of equilibrium counts for one, two, or n , according to its degree, the number so made up by the points or lines now considered will be less by one than the number of negatively charged bodies.

There are other points or lines of equilibrium which occur where the positive regions become separated from each other, and the negative region acquires periphraxy. The number of these, reckoned according to their degrees, is less by one than the number of positively charged bodies.

If we call a point or line of equilibrium positive when it is the meeting place of two or more positive regions, and negative when the regions which unite there are negative, then, if there are p bodies positively and n bodies negatively charged, the sum of the degrees of the positive points and lines of equilibrium will be $p - 1$, and that of the negative ones $n - 1$. The surface which surrounds the electrical system at an infinite distance from it is to be reckoned as a body whose charge is equal and opposite to the sum of the charges of the system.

But, besides this definite number of points and lines of equilibrium arising from the junction of different regions, there may be others, of which we can only affirm that their number must be even. For if, as any one of the negative regions expands, it becomes a cyclic region, and it may acquire, by repeatedly meeting itself, any number of degrees of cyclosis, each of which corresponds to the point or line of equilibrium at which the cyclosis was established. As the negative region continues to expand till it fills all space, it loses every degree of cyclosis it has acquired, and becomes at last acyclic. Hence there is a set of points or lines of equilibrium at which cyclosis is lost, and these are equal in number of degrees to those at which it is acquired.

If the form of the charged bodies or conductors is arbitrary, we can only assert that the number of these additional points or lines is even, but if they are charged points or spherical conductors, the number arising in this way cannot exceed $(n - 1)(n - 2)$ where n is the number of bodies*.

*{I have not been able to find any place where this result is proved.}

We finish the paper by mentioning that the last remark was added by J. J. Thomson in 1891 while proof-reading the third (and the last) edition of Maxwell's book. Adding the above numbers of obligatory and additional critical points one arrives at Conjecture 1.3 which was the starting point of our paper.

References

1. A. A. AGRACHEV, D. PALLASCHKE and S. SCHOLTES, 'On Morse theory for piecewise smooth functions', *J. Dynam. Control Systems* 3 (1997) 449–469.
2. D. N. BERNSTEIN, 'The number of roots of a system of equations', *Funktsional. Anal. i Priložen.* 9 (1975) no. 3, 1–4.
3. H. EDELSBRUNNER, *Algorithms in combinatorial geometry* (Springer, Berlin, 1987).
4. C. F. GAUSS, *Collected works*, 1st edn (Teubner, Leipzig, 1870–1917) vol. 3, p. 112; vol. 8, p. 32; vol. 10, part 2, pp. 189–191.
5. J. E. GOODMAN and J. O'ROURKE (eds), *Handbook of discrete and computational geometry* (CRC Press, Boca Raton, FL, 1997).
6. A. A. JANUŠAUSKAS, 'Critical points of electrostatic potentials', *Differentsial'nye Uravneniya i Primenen — Trudy Sem. Processov Optimal. Upravleniya. I Sekciya* 1 (1971) 84–90.
7. A. G. KHOVANSKII, *Fewnomials*, Translations of Mathematical Monographs 88 (American Mathematical Society, Providence, RI, 1991).
8. T. KIANG, 'On the critical points of non-degenerate Newtonian potentials', *Amer. J. Math.* 54 (1932) 92–109.
9. V. V. KOZLOV, 'A problem of Kelvin', *J. Appl. Math. Mech.* 81 (1990) 133–135.
10. A. G. KUŠNIRENKO, 'Newton polyhedra and Bezout's theorem', *Funktsional. Anal. i Priložen.* 10 (1976) no. 3, 82–83.
11. M. MARDEN, *Geometry of polynomials* (American Mathematical Society, Providence, RI, 1949).
12. J. C. MAXWELL, *A treatise on electricity and magnetism*, vol. 1, republication of the 3rd revised edition (Dover, New York, 1954).
13. P. McMULLEN and G. C. SHEPHARD, *Convex polytopes and the upper bound conjecture* (Cambridge University Press, 1971).
14. M. MORSE and S. CAIRNS, *Critical point theory in global analysis and differential topology* (Academic Press, New York, 1969).
15. T. S. MOTZKIN and J. L. WALSH, 'Equilibrium of inverse-distance forces in 3 dimensions', *Pacific J. Math.* 44 (1973) 241–250.
16. P. ORLIK and H. TERAOKA, 'The number of critical points of a product of powers of linear functions', *Invent. Math.* 120 (1995) 1–14.
17. F. PREPARATA and M. SHAMOS, *Computational geometry: an introduction* (Springer, Berlin, 1985).
18. R. SEIDEL, 'The complexity of Voronoi diagrams in higher dimensions', *Proceedings of the Allerton Conference on Communication, Control, and Computing* 20 (University of Illinois at Urbana-Champaign, 1982) 94–95.
19. D. SIERSMA, 'Voronoi diagrams and Morse theory of the distance function', *Geometry in present day science* (ed. O. E. Barndorff-Nielsen and E. B. V. Jensen; World Scientific, Singapore, 1999) 187–208.
20. A. VARCHENKO, 'Critical points of the product of powers of linear functions and families of bases of singular vectors', *Compositio Math.* 97 (1995) 385–401.

A. Gabrielov
 Department of Mathematics
 Purdue University
 West Lafayette, IN 47907-1395
 USA

agabriel@math.purdue.edu

D. Novikov
 Department of Mathematics
 Weizmann Institute of Science
 Rehovot, 76100
 Israel

dmitry.novikov@weizmann.ac.il

B. Shapiro
 Department of Mathematics
 University of Stockholm
 Stockholm S-10691
 Sweden

shapiro@math.su.se