ON THE THEORY OF POTENTIALS IN LOCALLY COMPACT SPACES

BY

BENT FUGLEDE

Copenhagen

Contents

Index of terminology and notations

Introduction

Since the modern theory of potentials was initiated by the works of O. Frostman [18], M. Riesz [26], and De la Vallée-Poussin [29], the further development has to a large extent centered around the following essential points: (1)

A. The discovery by H. Cartan [10] of the fact that the space $\mathcal{E}^+$ of all positive measures μ of *finite energy*

$$\|\mu\|^2 = \iint |x-y|^{2-n}\, d\mu(x)\, d\mu(y)$$

with respect to the Newtonian kernel in R^n, $n > 2$, is complete in the *strong topology*, i.e. the topology defined by the distance $\|\mu - \nu\|$.

B. The systematic use by J. Deny [15] of the *Fourier transform* in the sense of L. Schwartz [27] for the study of distributions of finite energy with respect to a positive definite distribution kernel invariant under translations in R^n.

(1) As to these and further lines of research, see the expository article on modern potential theory by M. Brelot [7].

C. The study of potentials of measures on a locally compact space with respect to a *regular* kernel (i.e. a kernel satisfying the principle of continuity, cf. § 2.1). The main tool is here the *vague topology*, which is related to the classical notion of convergence for positive measures. See G. Choquet [12, 13], M. Kishi [19], N. Ninomiya [23], M. Ohtsuka [24, 25].

D. The theory of abstract *capacities* developed by G. Choquet [14].

The present exposition is devoted to the theory of potentials of measures on a locally compact space, in particular on a group, the main emphasis being placed upon the study of *capacity* and *capacitary distributions*. In the existing literature concerning this subject the more advanced part of the theory is based on the assumption that the kernel be regular (cf. the references under C above), and it is usually required that the kernel fulfil Frostman's *maximum principle*. Instead of making assumptions of this nature we have chosen to base our study on the strong topology as well as the vague topology. Two papers by H. Cartan [9], [10] have served as a guide. We shall investigate the case of a positive *definite* kernel possessing the following two properties: (i) the space $\mathcal{E}^+$ is strongly complete (cf. above under A), and (ii) the strong topology on $\mathcal{E}^+$ ist stronger (= finer) than the vague topology on $\mathcal{E}^+$. A positive definite kernel possessing these two properties will be called *perfect*. (1) It turns out that the desired type of results concerning capacitary distributions (associated with arbitrary sets) and capacitability of analytic sets can be obtained in a simple and natural way in case of a perfect kernel (or just a consistent kernel). (1) These results are of a global nature, unlike the corresponding results based on the maximum principle and the vague topology. This fact reflects the *global character* of the concept of a perfect or a consistent kernel (cf. § 7.3). Apart from its global character, the concept of a consistent kernel is, however, more general than that of a kernel fulfilling Frostman's maximum principle. This becomes clear if we consider the case of a *compact* space (whereby the global aspects disappear). Then any positive kernel k satisfying the maximum principle is positive definite (Ninomiya [23]) and regular (Choquet [12]), and hence it follows from the proof of a theorem due to M. Ohtsuka [24] that k is consistent (cf. § 3.4 of the present paper).

The fact that the Newtonian kernel (and, more generally, the classical Green's function) is perfect, was proved by H. Cartan [10]. It is also known that the kernels

(1) A perfect kernel is, in particular, *strictly* definite, that is, the energy of a measure $\mu \neq 0$ is > 0 if at all defined. This strict definiteness (the so-called principle of energy) is, however, of minor importance for the development of the theory, and we have therefore introduced the weaker concept of a *consistent* kernel (§ 3.3). Such a kernel is definite, but not necessarily strictly definite.

$|x-y|^{\alpha-n}$ of order α in R^n are perfect $(0<\alpha<n)$, whereas the maximum principle is fulfilled for $\alpha \leqslant 2$ only. This perfect character of the kernels of order α (among other kernels) was established first by J. Deny [15, 16]. (1) We present an alternative, more elementary and more direct proof based exclusively on M. Riesz' composition formula (cf. Theorem 7.4 and § 8.1). The first result in this direction is due to H. Cartan [9], who proved that the kernels of order α are K-perfect, that is, perfect when considered only on compact subsets of the space. Actually, any strictly positive definite convolution kernel has this latter property, as we shall show in § 7. Altogether it turns out that practically all the definite kernels usually met with in analysis are consistent (and hence perfect if they are strictly positive definite).

The contents of the present paper may be summarized as follows.

Chapter I is of a preparatory character, and the methods and most of the results are well known. After a brief survey over the relevant parts of the theory of measures and integration on a locally compact space X (§ 1) follows an exposition of the theory of capacity and of the capacitary distributions (2) on compact sets (§ 2). In this section the potential and energy of measures are formed with respect to an arbitrary kernel on X. (A *kernel* on X is defined as a lower semi-continuous function $k=k(x, y)$ on $X \times X$.) (3) The proofs are based on the facts that potential and energy are lower semi-continuous functions on the space of all positive measures on X with the vague topology, and that the class of all positive measures of total mass 1 supported by a compact set is compact in the vague topology.

In Chapter II the kernel k is supposed to be positive *definite*. (We usually omit the qualification "positive".) The class $\mathcal{E}$ of all measures μ (of variable sign) of finite energy

---

(1) In Deny's theory, referred to above under B, is contained that a wide class of positive definite convolution kernels on R^n have properties very similar to those of a perfect kernel in our sense (cf. Deny [15], Chap. I, 3), the sole modification being that the concept of energy had to be defined in a manner quite different from the classical definition as a Lebesgue integral adopted in the present paper. In a supplementary paper Deny [16] proved that this difficulty can be overcome under the additional assumption that the kernel be regular. Nevertheless, it seems fair to say that the methods of Fourier analysis are not really adequate in the finer study of capacity and capacitary distributions.

(2) If the kernel fulfills Frostman's maximum principle, the capacitary distributions are also called *equilibrium distributions* because their potential is constant in the set in question (except in some subset of zero capacity). If the kernel is strictly definite, there is just one capacitary distribution on a given set.

(3) The most important case is that of a positive kernel: $0 \leqslant k(x, y) \leqslant +\infty$. We also admit kernels of variable sign, but in that case we restrict the attention to potentials and energy of measures of (uniformly) compact supports.

$$k(\mu, \mu) = \iint k(x, y)\, d\mu(x)\, d\mu(y) < \infty$$

is a pre-Hilbert space with the energy norm $\|\mu\| = (k(\mu, \mu))^{\frac{1}{2}}$. The subset $\mathcal{E}^+$ formed by all *positive* measures in $\mathcal{E}$ is a convex cone. Among the definite kernels we single out those for which the two topologies (strong and vague) on $\mathcal{E}^+$ have the following property of consistency: If a strong Cauchy filter Φ converges vaguely to some measure μ_0, then $\Phi \to \mu_0$ strongly. A definite kernel with this property will be called *consistent*. It is easily shown that a kernel is perfect if and only if it is consistent and strictly definite. Certain sufficient conditions of a general nature are obtained, and it is shown that a kernel obtained by superposition of consistent kernels is consistent. Under the hypothesis that the kernel be consistent we proceed to introduce the *interior and exterior capacitary distributions* associated with an arbitrary set of finite interior, resp. exterior, capacity. (1) We follow the method indicated by H. Cartan [10], § 6, for the Newtonian kernel (cf. also Aronszajn & Smith [1] for the kernels of order α), but certain modifications are required under the present general circumstances. As a by-product we obtain the following property of the exterior capacity of arbitrary sets:

$$\operatorname{cap}^* A = \lim_n \operatorname{cap}^* A_n$$

for any increasing sequence of subsets $A_n \subset X$. This result is the key to an application of Choquet's theory referred to above under D, and we conclude that every K-analytic subset of X is capacitable, i.e. of equal interior and exterior capacity. (2) The chapter ends with a brief discussion of the theory of "balayage".

Chapter III is devoted to the particularly interesting case in which the space X is a locally compact topological *group* and the kernel a convolution kernel, i.e.

$$k(x, y) = k(xy^{-1}),$$

(1) In the study of exterior capacitary distributions we must impose upon the locally compact space X a certain restriction, e.g. that X be metrizable.

(2) For the Newtonian kernel, as well as for Green's function, this result was obtained by G. Choquet [14], Chap. II, by application particularly of Cartan's maximum principle (H. Cartan [10], § VI), which leads to the fundamental inequality

$$\operatorname{cap}(A \cup B) + \operatorname{cap}(A \cap B) \leqslant \operatorname{cap} A + \operatorname{cap} B$$

for arbitrary compact sets A and B. The method described above in the text was used first by Aronszajn and Smith [1] in case of the kernels of order α. Recently, M. Kishi [19] has established the capacitability of all relatively compact Borelian or K-analytic subsets of a locally compact space of which every compact subset is metrizable, the assumption on the kernel being closely related (in fact equivalent) to Frostman's maximum principle.

the "kernel function" $k = k(x)$ being a given lower semi-continuous function on the group X. It is shown that any definite convolution kernel is K-consistent (Theorem 7.2). The proof is based on the fact that the *energy function*

$$k_\mu = \check{\mu} * k * \mu$$

is bounded and uniformly continuous provided $\mu \in \mathcal{E}^+$ (Theorem 7.1). This result follows, in turn, from a lemma asserting that a positive definite, lower semi-continuous function f on a topological group is bounded and uniformly continuous provided f is finite at the identity. The property of a definite convolution kernel to be actually consistent (not only K-consistent) depends, therefore, on the behaviour of the kernel function $k(x)$ as x tends to infinity in X, cf. § 7.3. In case of an Abelian group we prove, in particular, that if $k = k(x)$ has the form

$$k = \check{h} * h,$$

where $h \geqslant 0$ is lower semi-continuous on X, then the convolution kernel $k(xy^{-1})$ is consistent, and $\mathcal{E}^+$ is complete (Theorem 7.4). Several important kernels, including those of order α, are of this form. Some types of kernels studied by K. Kunugui [20], N. Ninomiya [22], and T. Ugaheri [28] are investigated further (§ 8.2), and the paper closes with a number of examples serving to illustrate various points in the theory.

I. Basic Concepts of Potential Theory

1. Measures on locally compact spaces

The distributions of mass (or charge) to be considered in the present study are those which can be interpreted mathematically as real-valued *measures*, in particular positive measures. Referring to N. Bourbaki [4], [5] for an exposition of the theory of measures and integration on a locally compact Hausdorff space, we limit ourselves to listing (in § 1.1) those concepts which are especially relevant in potential theory, and to stating (in § 1.2) some further necessary results. As to the terminology and notations we generally follow Bourbaki.[1]

[1] Certain exceptions from this convention will be listed here. Our locally compact (Hausdorff) space is denoted by X, and the class of all continuous functions on X by $\mathfrak{C} = \mathfrak{C}(X)$. (When speaking of a continuous function, we generally understand that the values are *finite* real numbers. A lower semi-continuous function is allowed to take the value $+\infty$, but not $-\infty$.) The class of all continuous functions of compact support is denoted by $\mathfrak{C}_0 = \mathfrak{C}_0(X)$. For any class $\mathfrak{F}$ of functions, $\mathfrak{F}^+$ denotes the class of all positive functions ($f \geqslant 0$) from $\mathfrak{F}$. The support of a function, or a measure, is denoted by $S(f)$, resp. $S(\mu)$. A set $A \subset X$ is said to be of class K_σ if A may be represented as the union of some sequence of compact subsets of X.

1.1. *Principal notions.* The class of all (Radon) measures on a locally compact (Hausdorff) space X is denoted by $\mathfrak{M} = \mathfrak{M}(X)$. On this linear space, the semi-norms $\mu \to \left|\int f \, d\mu\right|$, $f \in C_0$, define the so-called *vague topology*. A filter Φ on $\mathfrak{M}$ (cf. Bourbaki [2], Chap. I, § 5) converges vaguely to $\mu_0 \in \mathfrak{M}$ if

$$\lim_{\mu} \int f \, d\mu = \int f \, d\mu_0 \quad \text{along } \Phi \text{ for every } f \in C_0.$$

A subclass $\mathcal{B}$ of $\mathfrak{M}$ is called vaguely bounded if each of the above semi-norms remains bounded on $\mathcal{B}$. Any vaguely bounded subclass of $\mathfrak{M}$ is *relatively compact* in the vague topology (Bourbaki [4], Chap. III, § 2, N° 7). The converse is obvious since the semi-norms are continuous. Particularly useful is the induced vague topology on $\mathfrak{M}^+ = \mathfrak{M}^+(X)$, the class of all positive measures. The space $\mathfrak{M}^+$ is vaguely complete, and hence closed in $\mathfrak{M}$. More generally, the class $\mathfrak{M}_F^+$ of all positive measures supported by a given closed set F is a closed convex cone in $\mathfrak{M}$.

The *integral* (strictly speaking: upper integral) of a *lower semi-continuous* function $g \geq 0$ with respect to a measure $\mu \geq 0$ is defined by

$$\int g \, d\mu = \sup_{f \in C_0^+, f \leq g} \int f \, d\mu.$$

This integral is additive and positive homogeneous in g and μ. Clearly, the mapping $\mu \to \int g \, d\mu$, $\mu \in \mathfrak{M}^+$, is lower semi-continuous in the vague topology on $\mathfrak{M}^+$ for fixed lower semi-continuous $g \geq 0$.(¹) Since the characteristic function φ_G associated with an *open* set $G \subset X$ is lower semi-continuous, we may define the *measure* of G by $\mu(G) = \int \varphi_G \, d\mu$.

Subsequently one introduces the class $\mathcal{L}^1(\mu)$ of *μ-integrable* functions f with values $-\infty \leq f(x) \leq +\infty$ (Bourbaki [4], Chap. IV, §§ 3, 4). The integral of f with re-

(¹) Under the additional assumption that $\mu \in \mathfrak{M}^+$ be of compact support, one obtains a definition of $\int g \, d\mu$ for any lower semi-continuous g (whether positive or not) by replacing the class C_0^+ by the class C_0 in the above definition. It is easily verified that this integral is additive und positive homogeneous with respect to the two variables g and μ. Denoting by $c \geq 0$ a constant such that $g(x) \geq -c$ everywhere in $S(\mu)$, we get

$$\int g \, d\mu = \int (g + c) \, d\mu - c \int d\mu.$$

If X is compact, the mapping $\mu \to \int g \, d\mu$ is lower semi-continuous on $\mathfrak{M}^+$ for any given lower semi-continuous function g on X.

spect to μ is denoted by $\int f d\mu$. (1) Taking $f = \varphi_A =$ the characteristic function associated with a set $A \subset X$, we obtain the class of μ-integrable sets. The measure of such a set is defined by $\mu(A) = \int \varphi_A d\mu$. Any compact set is integrable. Next, one defines the local concepts of a *μ-measurable* function and a μ-measurable set (Bourbaki [4], Chap. IV, § 5).

For an arbitrary set $A \subset X$ the *exterior* and the *interior measure* of A are determined by

$$\mu^*(A) = \inf_{G \supset A} \mu(G) \quad \text{and} \quad \mu_*(A) = \sup_{K \subset A} \mu(K), \tag{1}$$

respectively. (The letter G refers to open sets and K to compact sets.) The equation $\mu^*(A) = \mu_*(A)$ subsists notably if A is open, or if A is μ-measurable and contained in the union of some sequence of μ-integrable sets. A set $N \subset X$ is called *μ-negligible* if $\mu^*(N) = 0$, and locally μ-negligible if $N \cap K$ is μ-negligible for every compact set K. These concepts lead to the notions *μ-almost everywhere* (μ-a.e.) and *locally μ-almost everywhere*, respectively.

The *trace* μ_A of a measure $\mu \geq 0$ on a μ-measurable set $A \subset X$ is defined by $\mu_A = \varphi_A \cdot \mu$, i.e.,

$$\int f d\mu_A = \int f \cdot \varphi_A d\mu, \qquad f \in C_0^+. \tag{2}$$

(Observe that $f\varphi_A$ is μ-integrable.) The total mass of μ_A is $\mu_A(X) = \mu_*(A)$ (cf. Lemma 1.2.2 below). A measure $\mu \geq 0$ is said to be *concentrated* on a set A if the complement $\complement A$ is locally μ-negligible; or, equivalently, if A is μ-measurable and $\mu = \mu_A$. It follows that $\mu(X) = \mu_*(A)$. (If A is closed, or if for instance $\mu^*(X) < +\infty$, then $\mu^*(\complement A) = \mu_*(\complement A) = 0$, that is, $\complement A$ is μ-negligible. A measure μ is, therefore, concentrated on a *closed* set A if and only if μ is *supported* by A, which means that $S(\mu) \subset A$.) For any measure $\mu \geq 0$ and any μ-measurable set A, μ_A is concentrated on A. We denote by $\mathfrak{M}_A^+$ the class of all positive measures concentrated on a given set $A \subset X$.

Using the canonical decomposition $\mu = \mu^+ - \mu^-$, one defines measurability and integrability with respect to a measure μ of *variable sign* by requiring measurability and integrability with respect to μ^+ and μ^-, or, equivalently, with respect to $|\mu| = \mu^+ + \mu^-$. For any μ-integrable function f one defines

(1) In particular, a lower semi-continuous function g (≥ 0 unless $S(\mu)$ is compact) is integrable with respect to a measure $\mu \geq 0$ if and only if $\int g d\mu$, as defined above, is finite. In the affirmative case the two notions of integral coincide.

$$\int f\,d\mu = \int f\,d\mu^+ - \int f\,d\mu^-.$$

The same formula serves to define the integral of a lower semi-continuous function f provided $\int f\,d\mu^+$ and $\int f\,d\mu^-$ are not both infinite; moreover, it is assumed either that $f \geqslant 0$ or that $S(\mu)$ is compact (cf. note (1), p. 145).

The *product* $\mu \otimes \nu$ of two measures μ and ν on the locally compact spaces X and Y, respectively, is defined as the unique measure on the product space $X \times Y$ such that

$$\int (\varphi \otimes \psi)\,d(\mu \otimes \nu) = \int \varphi\,d\mu \cdot \int \psi\,d\nu$$

for every $\varphi \in \mathfrak{C}_0(X)$, $\psi \in \mathfrak{C}_0(Y)$. Here $\varphi \otimes \psi$ denotes the function $\varphi(x)\cdot\psi(y)$ of class $\mathfrak{C}_0(X \times Y)$. When integrating with respect to $\mu \otimes \nu$ one may write $\iint f(x, y)\,d\mu(x)\,d\nu(y)$ instead of $\int f\,d(\mu \otimes \nu)$. The following two instances of Fubini's theorem

$$\iint f(x, y)\,d\mu(x)\,d\nu(y) = \int d\mu(x) \int f(x, y)\,d\nu(y) = \int d\nu(y) \int f(x, y)\,d\mu(x) \qquad (3)$$

will be used repeatedly in the sequel:

(i) If f is *integrable* with respect to $\mu \otimes \nu$, i.e. $f \in \mathfrak{L}^1(\mu \otimes \nu)$, then the interior integrals on the right represent functions of class $\mathfrak{L}^1(\mu)$ and $\mathfrak{L}^1(\nu)$, respectively, defined and finite almost everywhere, and (3) subsists.

(ii) If f is *lower semi-continuous* (and $\geqslant 0$ unless μ and ν have compact supports), then (3) holds provided the integral with respect to $\mu \otimes \nu$ is defined. (Cf. above. See also Lemma 1.2.6 below.) This is always the case if $\mu \geqslant 0$ and $\nu \geqslant 0$, in which case the interior integrals on the right represent lower semi-continuous functions of x and y, respectively.

As to the proofs, see Bourbaki [5], § 8, N° 1, for the case of positive measures; and apply Lemmas 1.2.3 and 1.2.6 below in the general case.

1.2. *Supplementary results.* In order to save space we omit the proofs of the following lemmas.

LEMMA 1.2.1. *If a locally compact space X is metrizable and of class K_σ, then $\mathfrak{M}^+(X)$ satisfies the first axiom of countability.*

In view of this lemma the use of filters may often be avoided in the sequel if one assumes that the locally compact space X is metrizable and of class K_σ (or, equivalently, that X satisfies the second axiom of countability, cf. Bourbaki [3], § 2, N° 9). In case of such a space X, a positive measure μ adheres to a subset $S \subset \mathfrak{M}^+(X)$ if and only if S contains a sequence converging to μ. In particular, S is closed if and only if S is sequentially closed. Similarly, any bounded sequence in $\mathfrak{M}^+$ contains a convergent subsequence.

We return to the study of measures on an arbitrary locally compact space X.

LEMMA 1.2.2. *For any lower semi-continuous function $g \geqslant 0$, any measure $\mu \geqslant 0$, and any μ-measurable set $A \subset X$,*

$$\int g\,d\mu_A = \sup_{K \subset A} \int g\,d\mu_K \quad (K \text{ compact}).$$

This is a generalization of the second part of (1), § 1.1, which corresponds to the case $g = 1$. The following lemma deals with measures of arbitrary sign.

LEMMA 1.2.3. *Suppose g is lower semi-continuous* (and $\geqslant 0$ unless μ and ν have compact supports). *The identity*

$$\int g\,d(a\mu + b\nu) = a\int g\,d\mu + b\int g\,d\nu$$

subsists in the sense that the integral on the left is defined whenever the two integrals on the right are both defined and the linear combination is meaningful.

The remaining three lemmas are concerned with the product of two measures on two locally compact spaces X and Y, respectively.

LEMMA 1.2.4. *The mapping $(\mu, \nu) \to \mu \otimes \nu$ of $\mathfrak{M}^+(X) \times \mathfrak{M}^+(Y)$ into $\mathfrak{M}^+(X \times Y)$ is continuous.*

(Cf. Bourbaki [4], exerc. 5, p. 100). The corresponding assertion concerning measures of arbitrary sign would be false.

LEMMA 1.2.5. *If $A \subset X$ and $B \subset Y$ are measurable with respect to $\mu \in \mathfrak{M}^+(X)$ and $\nu \in \mathfrak{M}^+(Y)$, respectively, then the trace of $\mu \otimes \nu$ upon $A \times B$ equals $\mu_A \otimes \nu_B$.*

In particular, $\mu \otimes \nu$ is concentrated on $A \times B$ if μ is concentrated on A and ν on B.

LEMMA 1.2.6. *The following identities hold for any two measures $\mu \in \mathfrak{M}(X)$ and $\nu \in \mathfrak{M}(Y)$:*

$$S(\mu \otimes \nu) = S(\mu) \times S(\nu),$$

$$|\mu \otimes \nu| = |\mu| \otimes |\nu|,$$

$$(\mu \otimes \nu)^+ = (\mu^+ \otimes \nu^+) + (\mu^- \otimes \nu^-),$$

$$(\mu \otimes \nu)^- = (\mu^+ \otimes \nu^-) + (\mu^- \otimes \nu^+).$$

2. Kernel, potential, energy, capacity

2.1. *Definitions.* By a *kernel* on a locally compact (Hausdorff) space X is meant a lower semi-continuous function $k = k(x, y)$, defined everywhere in $X \times X$, with values

$$-\infty < k(x, y) \leqslant +\infty.$$

A kernel is called *positive* if $k(x, y) \geqslant 0$ for every pair (x, y), and *strictly positive* if, in addition, $k(x, x) > 0$ for every $x \in X$. A kernel k is called *symmetric* if $k(x, y) = k(y, x)$. For any kernel k, the functions $k(x, \cdot)$ and $k(\cdot, y)$ are lower semi-continuous in X for fixed x and y, respectively.

We proceed to define potential and energy of measures with respect to a given kernel k. In order to avoid certain difficulties *we shall always assume that the measures in question have compact supports, except in the case of a positive kernel,* where arbitrary measures are admitted. The *potential* of a measure μ on X at a point $x \in X$ is then defined by

$$k(x, \mu) = \int k(x, y)\, d\mu(y) = k(x, \mu^+) - k(x, \mu^-)$$

provided $k(x, \mu^+)$ and $k(x, \mu^-)$ are not both infinite. In particular, the potential of a positive measure is defined everywhere and represents a lower semi-continuous function on X. We shall sometimes use the notation

$$k(A, \mu) = \sup_{x \in A} k(x, \mu) \qquad (A \subset X;\ \mu \geqslant 0).$$

The *mutual energy* of two measures μ and ν (of compact supports unless $k \geqslant 0$) is defined by

$$k(\mu, \nu) = \iint k(x, y)\, d\mu(x)\, d\nu(y) = k(\mu^+, \nu^+) + k(\mu^-, \nu^-) - k(\mu^+, \nu^-) - k(\mu^-, \nu^+) \quad (1)$$

provided $k(\mu^+, \nu^+) + k(\mu^-, \nu^-)$ or $k(\mu^+, \nu^-) + k(\mu^-, \nu^+)$ is finite (cf. Lemma 1.2.6); thus in particular if $\mu \geqslant 0$ and $\nu \geqslant 0$. From Fubini's theorem follows that

$$k(\mu, \nu) = \int k(x, \nu)\, d\mu(x) = \int k(\mu, y)\, d\nu(y)$$

whenever $k(\mu, \nu)$ is defined. For $\nu = \mu$ we obtain the *energy* of μ:

$$k(\mu, \mu) = \iint k(x, y)\, d\mu(x)\, d\mu(y) = \int k(x, \mu)\, d\mu(x).$$

If the kernel k is *symmetric,* we have the law of reciprocity

$$\int k(x, \mu)\, d\nu(x) = \int k(x, \nu)\, d\mu(x),$$

valid at least when $k(\mu, \nu)$ is defined.

The following three functions defined for positive measures μ (of compact support unless $k \geqslant 0$) present a certain similarity, and each of them gives rise to a concept of capacity (cf. § 2.3):

$$U(\mu) = k(X, \mu) \quad = \sup_{x \in X} k(x, \mu)$$

$$V(\mu) = k(S(\mu), \mu) = \sup_{x \in S(\mu)} k(x, \mu)$$

$$W(\mu) = k(\mu, \mu) \quad = \int k(x, \mu)\, d\mu.$$

Clearly, $$-\infty < W(\mu) \leqslant V(\mu) \cdot \mu(X); \qquad V(\mu) \leqslant U(\mu).$$

A kernel is said to satisfy Frostman's *maximum principle* if $U(\mu) = V(\mu)$ for every $\mu \in \mathfrak{M}^+$ of compact support. (1) (If a positive kernel satisfies this maximum principle, then $U(\mu) = V(\mu)$ for *every* $\mu \in \mathfrak{M}^+$.)

A kernel k will be called *pseudo-positive* if $W(\mu) \geqslant 0$ for every $\mu \in \mathfrak{M}^+$ of compact support; and *strictly pseudo-positive* if $W(\mu) > 0$ for every $\mu \in \mathfrak{M}^+$, $\mu \neq 0$, of compact support. Any positive kernel is pseudo-positive, and it is strictly pseudo-positive if and only if it is strictly positive. (In fact, if $k(\mu, \mu) = 0$, the open set of pairs (x, y) such that $k(x, y) > 0$ does not meet the support $S(\mu) \times S(\mu)$ of $\mu \otimes \mu$. Hence $k(x, x) = 0$ for every $x \in S(\mu)$. The converse statement is verified by taking $\mu = \varepsilon_x$ (= the mass $+1$ placed at the point x.)

A kernel k is called *regular* if it satisfies the *principle of continuity*, i.e. if one can conclude that the potential $k(x, \mu)$ of a measure $\mu \geqslant 0$ of compact support is continuous throughout X when it is known that the restriction of $k(x, \mu)$ to $S(\mu)$ is continuous. (2) As to the study of potentials with respect to a regular kernel, or a kernel satisfying Frostman's maximum principle, see for instance the literature referred to in the introduction (under C). In the present study we shall generally not make assumptions of this nature (cf., however, Theorems 3.4.1, 3.4.2, and 7.3).

---

(1) The fact that the Newtonian kernel satisfies this maximum principle was proved by M. A. J. Maria [21]. The kernels of orders $\alpha \leqslant 2$ have the same property, as shown by Frostman [18], p. 68.

(2) The regularity of the Newtonian kernel was proved by G. C. Evans [17], p. 238, and by F. Vasilesco [31]. For the kernels of order α, $0 < \alpha < n$, see Frostman [18], p. 26. More generally, S. Kametani has established the regularity of any kernel on R^n which is a continuous, decreasing, and positive function of $|x - y|$ (cf. K. Kunugui [20], p. 78). A further sufficient condition for regularity is found in H. Cartan & J. Deny [11], §§ 6, 7.

In the sequel we shall concentrate on either of the following two cases:

I: The kernel k is positive: $k(x, y) \geqslant 0$.

II: The space X is compact.

The remaining case of a kernel of variable sign on a locally compact, non compact space presents certain difficulties unless the attention is limited (as described above) to measures supported by some (fixed) compact subset $K \subset X$ (cf. § 5.3). This limitation actually amounts to replacing X by the compact space K, and k by its restriction to $K \times K$, so that one is back in Case II. Throughout the rest of Chapter I and the whole of Chapter II (except for § 5.3) we shall therefore always assume that one (or both) of the above cases I or II occurs. This general hypothesis will usually not be repeated. Case II can mostly be reduced to Case I simply by replacing the kernel k by the positive kernel k' obtained by adding to k a suitable constant $c \geqslant 0$:

$$k'(x, y) = k(x, y) + c \geqslant 0.$$

This is always possible since a lower semi-continuous function is bounded from below on a compact space.

A kernel k is called *definite* (= positive definite) if it is symmetric and if the energy $k(\mu, \mu)$ is $\geqslant 0$ whenever defined; and *strictly definite* if, in addition, $k(\mu, \mu) = 0$ implies $\mu = 0$. Thus a symmetric kernel is definite if and only if

$$k(\mu^+, \mu^+) + k(\mu^-, \mu^-) \geqslant 2k(\mu^+, \mu^-)$$

for every measure μ. Any definite kernel is pseudo-positive, and any strictly definite kernel is strictly pseudo-positive. Chapters II and III are devoted to the study of potentials with respect to a definite kernel.

In the remaining part of Chapter I, only *positive measures* will be considered, and the space $\mathfrak{M}^+$ of all such measures will be thought of as a Hausdorff space with the vague topology (§ 1.1).

2.2. *Potential and energy of positive measures.*

LEMMA 2.2.1. *The following five functions are lower semi-continuous:*

(a) $k(\mu, \nu) = \iint k(x, y)\, d\mu(x)\, d\nu(y)$ *on* $\mathfrak{M}^+ \times \mathfrak{M}^+$.

(b) $k(x, \mu) = \int k(x, y)\, d\mu(y)$ *on* $X \times \mathfrak{M}^+$.

(c) $U(\mu) = k(X, \mu)$ *on* $\mathfrak{M}^+$.

(d) $V(\mu) = k(S(\mu), \mu)$ *on* $\mathfrak{M}^+$.

(e) $W(\mu) = k(\mu, \mu)$ *on* $\mathfrak{M}^+$.

Proof. Ad (a): consequence of Lemma 1.2.4 since the function $\int k\,d\lambda$ of $\lambda \in \mathfrak{M}^+(X\times X)$ is lower semi-continuous (§ 1.1). Ad (b): consequence of (a) because the mapping $x \to \varepsilon_x$ of X into $\mathfrak{M}^+$ is continuous and $k(x, \mu) = k(\varepsilon_x, \mu)$ when ε_x denotes the mass $+1$ placed at the point $x \in X$. Ad (c) and (d): consequences of (b). We show this in the case of the function $V(\mu)$. Let $\mu_0 \in \mathfrak{M}^+$, $t < V(\mu_0)$. Then $k(x_0, \mu_0) > t$ for some $x_0 \in S(\mu_0)$. In view of (b) there are neighbourhoods A of x_0 in X and B of μ_0 in $\mathfrak{M}^+$ such that

$$k(x, \mu) > t \qquad \text{for } x \in A,\ \mu \in B. \tag{1}$$

Since $x_0 \in S(\mu_0)$, there is a function $f \in C_0^+$ with $S(f) \subset A$ such that $\int f\,d\mu_0 \neq 0$, and hence $\int f\,d\mu \neq 0$ for every μ in some neighbourhood B' of μ_0. This implies that $S(\mu)$ has some point x in common with A when $\mu \in B'$. Using this point x in (1), we conclude that $V(\mu) \geq k(x, \mu) > t$ for every $\mu \in B \cap B'$. Ad (e): consequence of (a).

LEMMA 2.2.2. *If a positive measure μ is concentrated on some set $A \subset X$, then*

$$U(\mu) = \lim_{K\uparrow A} U(\mu_K);\quad V(\mu) = \lim_{K\uparrow A} V(\mu_K);\quad W(\mu) = \lim_{K\uparrow A} W(\mu_K),$$

where μ_K denotes the trace of μ upon K, and K ranges over the increasing filtering family of all compact subsets of A.[1]

Proof. According to Lemma 1.2.2. with $g \in C_0^+$, $\mu_K \to \mu_A = \mu$ vaguely as $K \uparrow A$. Hence it follows from the preceding lemma that

$$U(\mu) \leq \lim_{K\uparrow A} U(\mu_K) \leq U(\mu)$$

(in Case I), and similarly with V or W in place of U. Case II (X compact) follows from Case I in the usual way.

Remark. If $\mu \geq 0$ and $\nu \geq 0$ are concentrated on $A \subset X$ and $B \subset X$, respectively, it follows in the same way that

$$k(\mu, \nu) = \lim_{H\uparrow A} k(\mu_H, \nu) = \lim_{K\uparrow B} k(\mu, \nu_K) = \lim_{H\uparrow A,\, K\uparrow B} k(\mu_H, \nu_K),$$

where H and K denote arbitrary compact subsets of A and B, respectively.

---

(1) As to the concept of a filtering family (generalizing that of a monotone sequence), we refer to Bourbaki [2], Chap. I, § 5, N° 4. In case I ($k \geq 0$) the sign "lim" may be replaced by supremum over all compact subsets $K \subset A$.

2.3. *Capacities associated with a kernel.* Given a kernel $k = k(x, y)$ on X, we derive from each of the functions $U(\mu)$, $V(\mu)$, $W(\mu)$ a set function by defining, for any set $A \subset X$,

$$u(A) = \inf U(\mu); \;\; v(A) = \inf V(\mu); \;\; w(A) = \inf W(\mu), \tag{1}$$

where μ ranges over the class of all positive measures concentrated on A and of total mass $\mu(X) = \mu(A) = 1$. (We interpret these infima as $+\infty$ if A is void.) The result would be the same if $S(\mu)$ were required to be compact and contained in A; this follows easily from Lemma 2.2.2 and the second relation (1), § 1.1. Hence

$$u(A) = \inf u(K); \;\; v(A) = \inf v(K); \;\; w(A) = \inf w(K), \tag{2}$$

where K ranges over the class of all compact subsets of A. Clearly, each of the three set functions is *decreasing* and attains its minimum at $A = X$. Moreover,

$$+\infty \geqslant u(A) \geqslant v(A) \geqslant w(A) > -\infty.$$

If the kernel satisfies Frostman's maximum principle, $u(A) = v(A)$ for every set A. It is well known, and will be shown in § 2.4, that $v(A) = w(A)$ for every set A if k is symmetric.

To each of the functions u, v, w corresponds an "exterior" set function defined as follows for arbitrary $A \subset X$:

$$u^*(A) = \sup u(G); \;\; v^*(A) = \sup v(G); \;\; w^*(A) = \sup w(G), \tag{3}$$

where G ranges over the class of all open sets containing A. The relations $u^*(A) = u(A)$, etc., hold for any open set A. It will be shown presently that they hold likewise if A is compact (Lemma 2.3.4).

The sets $N \subset X$ such that $w(N) = +\infty$ or $w^*(N) = +\infty$ play an important role as negligible sets. Observe that each of these two classes of negligible sets remains unchanged if the kernel k is replaced by $k + c$ for some constant c. In the study of the two types of negligible sets it suffices, therefore, to consider case I ($k \geqslant 0$). A proposition involving a variable point $x \in A$ (where A denotes a given subset of X) is said to subsist *nearly everywhere* (n.e.) in A if $w(N) = +\infty$, N being the set of points of A for which the proposition fails to hold. Similarly, the proposition is said to hold *quasi-everywhere* (q.e.) in A if $w^*(N) = +\infty$. The following lemma is easily verified, for instance in the succession (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) $\Rightarrow$ (iv) $\Rightarrow$ (v) $\Rightarrow$ (i):

LEMMA 2.3.1. *Let $N \subset X$. The following five conditions are equivalent:*

(i) $w(N) = +\infty$.

(ii) $w(K) = +\infty$ for every compact set $K \subset N$.

(iii) $\mu_*(N) = 0$ for every positive measure μ of finite energy on X.

(iv) $\mu = 0$ is the only positive measure of finite energy concentrated on N.

(v) $\mu = 0$ is the only positive measure of finite energy supported by some compact subset of N.

We denote in the sequel by $\mathfrak{A}$ the class of all sets which are measurable with respect to every measure on X. It follows from condition (iii) of the above lemma that *the union of any denumerable family of sets of class $\mathfrak{A}$ with $w = +\infty$ is a set with $w = +\infty$*. The following lemma will be used later.

LEMMA 2.3.2. *Let $\mu \in \mathfrak{M}^+$, $A \subset X$, and $0 \leqslant t \leqslant +\infty$. The following two propositions are equivalent:*

(a) $k(x, \mu) \geqslant t$ for nearly every $x \in A$.

(a′) $k(\nu, \mu) \geqslant t \cdot \nu(X)$ for every positive measure ν of finite energy and supported by some compact subset of A.

In proving that (a) implies (a′), we may suppose $k(\nu, \mu) < \infty$, i.e. $k(x, \mu)$ is ν-integrable. Writing $N = \{x \in S(\nu) : k(x, \mu) < t\}$, we have $w(N) = +\infty$, and hence $\nu_*(N) = 0$. Since N is ν-measurable and contained in the compact support of ν, we conclude that $\nu^*(N) = 0$, and hence

$$k(\nu, \mu) = \int_{S(\nu)} k(x, \mu)\, d\nu \geqslant t \cdot \int_{S(\nu)} d\nu = t \cdot \nu(X).$$

Next, suppose (a) is not fulfilled, and write $B = \{x \in A : k(x, \mu) < t\}$. Then $w(\dot{B}) \neq +\infty$, and hence there is, in view of Lemma 2.3.1, a positive measure $\nu \neq 0$ of finite energy supported by some compact set $K \subset B$. Since $k(x, \mu) < t$ in K, we obtain $k(\nu, \mu) < t \cdot \nu(X)$ in contradiction with (a′).

THEOREM 2.3. *For any non-empty compact set $K \subset X$, each of the three infima (1), with $A = K$, is an actual minimum. The minimizing measures constitute a compact subclass of $\mathfrak{M}^+$.*

Proof. Choose a function $f \in C_0^+$ which equals 1 in K. Then $\mu(X) = \int f\, d\mu$ depends continuously on $\mu \in \mathfrak{M}_K^+$. The subset of $\mathfrak{M}^+$ over which μ ranges in the extremal problems (1) is therefore closed (in the vague topology). Being vaguely bounded, this set is compact (and non-empty when K is non-empty). Hence the theorem follows from Lemma 2.2.1. We denote the three classes of minimizing measures by $\mathfrak{U}_K$, $\mathfrak{V}_K$, and $\mathfrak{W}_K$, respectively.

LEMMA 2.3.3. *If A denotes the union of an increasing sequence of sets A_n of class $\mathfrak{A}$, then*

$$u(A) = \lim_{n\to\infty} u(A_n); \quad v(A) = \lim_{n\to\infty} v(A_n); \quad w(A) = \lim_{n\to\infty} w(A_n).$$

Proof. It suffices to consider case I ($k \geqslant 0$), and we may assume that A is non-void. Let $\mu \in \mathfrak{M}_A^+$, $\mu(X) = 1$, and denote by μ_n the trace of μ on the μ-measurable set A_n. Then $\mu_n(X) = \mu(A_n) \to \mu(A) = \mu(X) = 1$, and hence we may suppose all $\mu_n \neq 0$. Writing $\lambda_n = \mu_n/\mu(A_n)$, we have $\lambda_n \in \mathfrak{M}_{A_n}^+$ and $\lambda_n(X) = 1$. Hence

$$w(A_n) \leqslant W(\lambda_n) = W(\mu_n)/\mu(A_n)^2 \leqslant W(\mu)/\mu(A_n)^2.$$

Letting $n \to \infty$, we conclude that $\lim w(A_n) \leqslant W(\mu)$, which implies the non-trivial part of the limit relation for w. The proof is similar for the functions u and v.

Remark. For a *decreasing* sequence the corresponding limit relation would be false in general.(1) It does hold, however, in the case of *compact* sets; and in that case the sequence may even be replaced by an arbitrary decreasing *filtering family* $\mathfrak{F}$ of compact sets K. Writing $K_0 = \bigcap_{K \in \mathfrak{F}} K$, we propose to establish the following relations:

$$u(K_0) = \lim u(K); \quad v(K_0) = \lim v(K); \quad w(K_0) = \lim w(K) \tag{4}$$

as $K \to K_0$ through $\mathfrak{F}$. (Of course, the limit sign may be replaced by supremum over all $K \in \mathfrak{F}$.) Consider, e.g., the set function w, and choose for every $K \in \mathfrak{F}$ some minimizing measure $\mu_K \in \mathfrak{W}_K$ (i.e. $\mu_K \in \mathfrak{M}_K^+$, $\mu_K(X) = 1$, $W(\mu_K) = w(K)$. We disregard the trivial case where some K is empty). Since the measures μ_K, $K \in \mathfrak{F}$, belong to the bounded (i.e. relatively compact) class of all measures $\mu \in \mathfrak{M}^+$ with $\mu(X) \leqslant 1$, there exists at least one clusterpoint μ_0 of μ_K as $K \to K_0$ through $\mathfrak{F}$. By virtue of the lower semi-continuity of $W(\mu)$ and the fact that w is decreasing, we conclude that, as $K \to K_0$ through $\mathfrak{F}$,

$$W(\mu_0) \leqslant \lim w(K) \leqslant w(K_0). \tag{5}$$

Clearly μ_0 is supported by each $K \in \mathfrak{F}$ and hence by K_0. It will be shown presently

(1) As an example pertaining to the Newtonian kernel $|x-y|^{-1}$ in R^3, let A_n denote the open region between two concentric spheres, say $A_n = \{x \in R^3 : 1 - n^{-1} < |x| < 1\}$. Then $\bigcap_n A_n$ is void, but $w(A_n) = 1$.

that $\mu_0(X)=1$. Then it follows from (5) that $w(K_0)=W(\mu_0)$. This implies, in turn, the desired relation (4) and, moreover, that μ_0 *is a minimizing measure for the set* K_0.(1) To see that $\mu_0(X)=1$, we choose a function $f\in C_0^+$ so that $f(x)=1$ in some open set $G\supset K_0$. Now, G contains some set $H\in\mathfrak{F}$.(2) For every $K\in\mathfrak{F}$ such that $K\subset H$ we have

$$\int f\,d\mu_K=\mu_K(X)=1,$$

and hence
$$\mu_0(X)=\int f\,d\mu_0=\lim\int f\,d\mu_K=1.$$

LEMMA 2.3.4. *For any compact set* K, $u^*(K)=u(K)$, $v^*(K)=v(K)$, *and* $w^*(K)=w(K)$.

This follows from the above remark applied to the filtering family of all compact neighbourhoods H of K. It is well known that the intersection of this family is K. Hence there corresponds to every number $t<w(K)$ a compact neighbourhood H of K such that $w(H)>t$. If G denotes the interior of H, we have $G\supset K$, and hence

$$w^*(K)\geqslant w(G)\geqslant w(H)>t.$$

Consequently, $w^*(K)\geqslant w(K)$, q.e.d.

This lemma, or the remark on which it was based, asserts that the decreasing set functions u, v, and w are *continuous from the outside* when considered on compact sets. Now, let $\theta=\theta(t)$ denote a *decreasing* function of the real variable t, mapping the interval $w(X)\leqslant t\leqslant+\infty$ (in case of the set function w) in a continuous way into the extended real line. The set function $\gamma(K)=\theta(w(K))$, considered on the class of

(1) The part of this remark which concerns the minimizing measures can be formulated in a slightly stronger way in which the arbitrary choice of a minimizing measure for each set $K\in\mathfrak{F}$ is avoided. We associate with every set $H\in\mathfrak{F}$ the "section"

$$\mathfrak{S}(H)=\bigcup_{K\subset H}\mathfrak{W}_K\qquad(H\in\mathfrak{F},\ K\in\mathfrak{F})$$

fonsisting of all minimizing measures for all subsets $K\subset H$, $K\in\mathfrak{F}$. These sections $\mathfrak{S}(H)$ generate a *ilter* Φ on a relatively compact subset of $\mathfrak{M}^+$ (because $\mu(X)=1$), and the proof above shows that the vague adherence of Φ is contained in $\mathfrak{W}_{K_0}$.

(2) In fact, the decreasing filtering family formed by the compact sets $K\cap\complement G$, $K\in\mathfrak{F}$, has a void intersection. According to the "finite intersection principle", there is a finite family $\{K_i\}_{i=1}^n$, $K_i\in\mathfrak{F}$, such that the intersection

$$\Big(\bigcap_{i=1}^n K_i\Big)\cap\complement G=\bigcap_{i=1}^n(K_i\cap\complement G)$$

is void. Since $\mathfrak{F}$ is filtering, there is a set $H\in\mathfrak{F}$ such that $H\subset\bigcap_{i=1}^n K_i$, and hence $H\subset G$.

all compact sets, is then a *capacity* in the sense of Choquet [14], § 15, (i.e. an increasing set function which is continuous from the outside). According to (2), resp. (3),

$$\gamma_*(A) = \theta(w(A)) \quad \text{and} \quad \gamma^*(A) = \theta(w^*(A))$$

are the corresponding interior, resp. exterior, capacities. A set A is called *capacitable* if $\gamma_*(A) = \gamma^*(A)$. For such a set we write simply $\gamma(A)$ instead of $\gamma_*(A)$ or $\gamma^*(A)$. Similarly, one may study the capacities $\alpha(K) = \theta(u(K))$ and $\beta(K) = \theta(v(K))$. As shown above, open sets and compact sets are capacitable. Results concerning capacitability of more general sets can be obtained by application of Choquet's theory, at least under suitable assumptions concerning the kernel k and the space X, cf. M. Kishi [19] in the case of a kernel $k \geqslant 0$ satisfying Frostman's maximum principle. In § 4 of the present study we obtain somewhat stronger results in case of a consistent kernel.

We shall now consider the following choice of the function θ:

$$\theta(t) = \frac{1}{t-a},$$

where a denotes a real constant $\leqslant w(X)$ (resp. $u(X)$ or $v(X)$). In particular, any number $a \leqslant \inf k(x, y)$ can be used. The most important case is $a = 0$, which leads to the Wiener capacity, cf. § 2.5.

LEMMA 2.3.5. *If a denotes a constant such that $k(x, y) \geqslant a$, then the interior capacity $(w(A) - a)^{-1}$ is countably subadditive on sets of class $\mathfrak{A}$, and the exterior capacity $(w^*(A) - a)^{-1}$ is countably subadditive on arbitrary sets. Similarly with w replaced by u or v.*

Proof. The kernel $k' = k - a$ is positive, and the corresponding function w' equals $w - a$. Hence the capacity $1/w'$ equals the capacity in question. It suffices, therefore, to prove the lemma for a *positive kernel* k (with $a = 0$). Consider a sequence of sets A_n with the union A. Our task is to prove, first, that

$$w(A)^{-1} \leqslant \sum_n w(A_n)^{-1} \quad \text{provided } A_n \in \mathfrak{A}. \tag{6}$$

Without loss of generality we may assume that the sets A_n are mutually disjoint and that A is not void. For any positive measure μ with compact support contained in A and with $\mu(X) = 1$, we denote the trace of μ upon A_n by μ_n and write $\lambda_n = \mu_n / \mu(A_n)$ for such values of n that $\mu(A_n) \neq 0$. Neglecting the remaining values of n, if any, we have

$$w(A_n) \leqslant k(\lambda_n, \lambda_n) = k(\mu_n, \mu_n)/\mu(A_n)^2,$$

and hence, by application of Cauchy's inequality,

$$\sum_n w(A_n)^{-1} \geqslant \sum_n \mu(A_n)^2/k(\mu_n, \mu_n) \geqslant (\sum_n \mu(A_n))^2/\sum_n k(\mu_n, \mu_n).$$

The resulting inequality holds *a fortiori* if the summations are extended over *all* indices n. Note that $\sum \mu(A_n) = \mu(A) = \mu(X) = 1$ because the sets A_n are μ-measurable and mutually disjoint. Hence (6) will follow if we can show that

$$\sum_n k(\mu_n, \mu_n) \leqslant k(\mu, \mu).$$

This inequality is easily derived from the corresponding inequality in which the kernel k is replaced by an arbitrary function $f \in C_0^+(X \times X)$ with $f \leqslant k$; and in that case we may apply (2), § 1.1 (with μ replaced by $\mu \otimes \mu$ and A by $A_n \times A_n$):

$$\sum_n \int f\, d(\mu_n \otimes \mu_n) = \sum_n \int_{A_n \times A_n} f\, d(\mu \otimes \mu) \leqslant \int f\, d(\mu \otimes \mu),$$

the sets $A_n \times A_n$ being mutually disjoint. Having thus established (6), we infer from the definition (3) of w^* that $1/w^*$ is countably subadditive on arbitrary sets.—The corresponding assertions concerning the capacities $1/u$ and $1/v$ may be verified similarly, but it is considerably simpler to make use of the following characterizations of $1/u$ and $1/v$, valid if $u(X) \geqslant 0$, resp. $v(X) \geqslant 0$; thus in particular if the kernel k is positive ($k \geqslant 0$) or pseudo-positive ($w(X) \geqslant 0$):

$$1/u(A) = \sup \lambda_*(A) \qquad (\lambda \in \mathfrak{M}_A^+;\ k(x, \lambda) \leqslant 1 \text{ everywhere})$$

$$1/v(A) = \sup \lambda_*(A) \qquad (\lambda \in \mathfrak{M}_A^+;\ k(x, \lambda) \leqslant 1 \text{ for } x \in S(\lambda)). \tag{7}$$

COROLLARY. Let k denote an arbitrary kernel ($\geqslant 0$ unless X is compact), and let N denote the union of a sequence of sets $N_n \subset X$. If $N_n \in \mathfrak{A}$ and $w(N_n) = +\infty$, then $w(N) = +\infty$. If $w^*(N_n) = +\infty$, then $w^*(N) = +\infty$. Likewise $w(A \cup N) = w(A)$ if $w(N) = +\infty$ and A, $N \in \mathfrak{A}$; $w^*(A \cup N) = w^*(A)$ if $w^*(N) = +\infty$. Similar statements apply to the set functions u and v.

Remark. In Lemma 2.3.5 (and its corollary above) the assumption that the sets A_n (resp. N_n) be of class $\mathfrak{A}$ (in case of the functions u, v, w) may be replaced by the slightly weaker hypothesis that $A_n = A_n' \cap B$, where $A_n' \in \mathfrak{A}$, whereas B is arbitrary. Writing $A' = \bigcup_n A_n'$, we have then $A = A' \cap B$. Since $S(\mu) \subset A \subset B$, the set B is μ-measurable, and so are, therefore, the sets A_n. Note that the restriction to positive kernels is indispensable in Lemma 2.3.5 (cf. Ex. 2, § 8.3).

2.4. *The case of a symmetric kernel.* When the kernel k is symmetric, the two set functions v and w are identical, and so are v^* and w^*. It suffices to prove that $v(K)=w(K)$ for every *compact* set K. Since we know that $v \geqslant w$, we may assume that $w(K) < +\infty$. We propose to verify, moreover, that the two minimum problems defining $v(K)$ and $w(K)$ (cf. Theorem 2.3):

$$v(K)=\min V(\mu); \qquad w(K)=\min W(\mu),$$

(where in both cases $\mu \in \mathfrak{M}_K^+$ and $\mu(K)=1$) have precisely the *same solutions*, i.e. $\mathfrak{V}_K=\mathfrak{W}_K$. In the theorem below we prove that $\mu \in \mathfrak{W}_K$ implies $V(\mu) \leqslant w(K)$ (property (b)). Since $w(K) \leqslant v(K)$, we infer that $\mu \in \mathfrak{V}_K$ and that, actually, $w(K)=v(K)$. Conversely, $\mu \in \mathfrak{V}_K$ implies $k(\mu, \mu) \leqslant V(\mu)=v(K)=w(K)$, which shows that $\mu \in \mathfrak{W}_K$. Consequently, $\mathfrak{V}_K=\mathfrak{W}_K$. The solutions $\mu \in \mathfrak{W}_K$ of the second (and hence of the first) minimum problem above will be called *capacitary distributions of unit mass* on K.

THEOREM 2.4. *Let k denote a symmetric kernel, and K a compact set such that $w(K) < +\infty$. The potential of any capacitary distribution μ of unit mass on K has the following properties:*

(a) $k(x, \mu) \geqslant w(K)$ nearly everywhere in K.

(b) $k(x, \mu) \leqslant w(K)$ everywhere in the support of μ.

(c) $k(x, \mu) = w(K)$ μ-almost everywhere in X.

Proof. We begin by establishing (a) in the equivalent integrated form (cf. Lemma 2.3.2):

(a′) $k(\nu, \mu) \geqslant w(K) \cdot \nu(X)$ for every $\nu \in \mathfrak{M}_K^+$ with $k(\nu, \nu) < +\infty$. (1)

It suffices to consider the case where $k(\nu, \mu) < +\infty$ and $\nu(X)=1$. Then $a\mu + b\nu$ is a "competing" measure for any choice of constants $a \geqslant 0$, $b \geqslant 0$, with $a+b=1$. Hence its energy

$$a^2 k(\mu, \mu) + 2ab \cdot k(\nu, \mu) + b^2 k(\nu, \nu)$$

attains its minimum at $a=1$, $b=0$. This implies (a′) when it is observed that

(1) The assumption $k(\nu, \nu) < +\infty$ is essential, not only for the proof, but for the validity of the result. Otherwise one could take $\nu=\varepsilon_x$ (= the mass $+1$ placed at an arbitrary point $x \in K$) and conclude that $k(x, \mu) \geqslant w(K)$ *everywhere* in K. This would be false even in case of the Newtonian kernel (unless the unbounded component of $\complement K$ is regular for Dirichlet's problem).

$k(\mu, \mu) = w(K)$. From (a) follows in view of Lemma 2.3.1 that $\mu_*(N) = 0$ when N denotes the set of points $x \in K$ such that $k(x, \mu) < w(K)$. Since N is measurable and μ finite, we conclude that $\mu^*(N) = 0$. Moreover, $\mu^*(\complement K) = 0$, and hence $k(x, \mu) \geqslant w(K)$ μ-a.e. in X. Integrating with respect to μ, we obtain

$$w(K) = k(\mu, \mu) = \int k(x, \mu)\, d\mu(x) \geqslant w(K) \cdot \mu(X) = w(K),$$

so that, actually, $k(x, \mu) = w(K)$ μ-a.e. Having thus obtained (c), we complete the proof by observing that (b) is equivalent to the following consequence of (c):

(b′) $\quad k(x, \mu) \leqslant w(K)$ μ-almost everywhere in X.

In fact, the set $G = \{x \in X : k(x, \mu) > w(K)\}$ is open, and hence $\mu(G) = 0$ if and only if $G \cap S(\mu)$ is void.

Remark 1. If the symmetric kernel k satisfies Frostman's *maximum principle*, then $U(\mu) = V(\mu)$, and hence we obtain the following stronger versions of (a) and (b):

(a_1) $\quad k(x, \mu) = w(K)$ nearly everywhere in K.

(b_1) $\quad k(x, \mu) \leqslant w(K)$ everywhere in X.

In view of (a_1) the capacitary distributions $\mu \in \mathfrak{W}_K$ $(= \mathfrak{U}_K = \mathfrak{V}_K)$ are then called *equilibrium distributions* (of unit mass) on K. Cf. Frostman [18], §§ 17, 31, for the kernels of order $\alpha \leqslant 2$. For *any* value of α, $0 < \alpha < n$, it was shown by Frostman that the inequality $k(x, \mu) \geqslant w(K)$ holds for every interior point x of K.

Remark 2. If k is *definite*, the class $\mathfrak{W}_K$ $(= \mathfrak{V}_K)$ of all capacitary distributions of unit mass on a compact set K with $w(K) < +\infty$ is *convex* (and compact) and consists of all competing measures $\mu \in \mathfrak{M}_K^+$, $\mu(X) = 1$, such that $k(x, \mu) \geqslant W(\mu)$ nearly everywhere in K; or equivalently

$$k(\nu, \mu) \geqslant k(\mu, \mu) \qquad \text{for every } \nu \in \mathfrak{M}_K^+ \text{ with } \nu(X) = 1 \text{ and } k(\nu, \nu) < \infty.$$

(In fact, this inequality implies, when applied to the capacitary distributions $\nu \in \mathfrak{W}_K$,

$$0 \leqslant k(\nu - \mu, \nu - \mu) = k(\nu, \nu) + k(\mu, \mu) - 2k(\nu, \mu) \leqslant k(\nu, \nu) - k(\mu, \mu).$$

Thus
$$k(\mu, \mu) \leqslant k(\nu, \nu) = w(K),$$

and consequently $\mu \in \mathfrak{W}_K$.) Observe also that $k(\nu - \mu, \nu - \mu) = 0$ for any two measures μ and ν of class $\mathfrak{W}_K$. This shows that, in case of a *strictly definite* kernel, there is just one capacitary distribution of unit mass on K.

2.5. *The case of a pseudo-positive kernel.* A kernel k is pseudo-positive (§ 2.1) if and only if $w(X) \geqslant 0$, or equivalently if $w(K) \geqslant 0$ for every compact set K. A kernel is strictly pseudo-positive if and only if $w(K) > 0$ for every compact set K.

LEMMA 2.5.1. *If the kernel k is strictly pseudo-positive, the set of all positive measures μ such that $k(\mu, \mu) \leqslant M$ is compact in the vague topology for any constant $M \geqslant 0$.*

Proof. The set $\mathcal{H} \subset \mathfrak{M}^+$ in which $k(\mu, \mu) \leqslant M$ is closed by virtue of Lemma 2.2.1. It remains to be proved that $\mathcal{H}$ is relatively compact, or equivalently that $\mathcal{H}$ is (vaguely) bounded:

$$\sup_{\mu \in \mathcal{H}} \int f \, d\mu < +\infty \qquad \text{for every } f \in C_0^+.$$

In Case I the hypothesis is: $k(x, y) \geqslant 0$ and $k(x, x) > 0$. In view of the lower semi-continuity of k, each point $x_0 \in X$ has a neighbourhood A such that $k(x, y) \geqslant a$ for $(x, y) \in A \times A$, $a > 0$ being a suitable constant. Let φ denote a function $\leqslant 1$ of class C_0^+ which equals 1 in some closed neighbourhood $B \subset A$ of x_0 and vanishes outside A. Then $a \cdot \varphi(x) \varphi(y) \leqslant k(x, y)$ for all x and y, and hence

$$a \cdot \left(\int \varphi \, d\mu\right)^2 \leqslant k(\mu, \mu) \leqslant M$$

for every $\mu \in \mathcal{H}$. Clearly, any function $f \in C_0^+$, say with $f \leqslant 1$, is dominated by some finite sum of functions φ obtained in this manner, and consequently $\int f \, d\mu$ remains bounded on $\mathcal{H}$. In Case II, the space X is compact, and hence $w(X) > 0$ (because k is strictly pseudo-positive). For reasons of homogeneity,

$$1/w(X) = \sup_{\mu} \mu(X)^2 / k(\mu, \mu) \qquad (\mu \in \mathfrak{M}^+, \ \mu \neq 0),$$

and hence $\mu(X)$ is bounded on $\mathcal{H}$, q.e.d.

If k is pseudo-positive, $1/w$ is a capacity called the *Wiener capacity.* (1) We

(1) The three capacities defined on compact sets K by

$$1/u(K) = \max \lambda(K) \qquad (\lambda \in \mathfrak{M}_K^+, \ U(\lambda) \leqslant 1),$$
$$1/v(K) = \max \lambda(K) \qquad (\lambda \in \mathfrak{M}_K^+, \ V(\lambda) \leqslant 1),$$
$$1/w(K) = \max \{2\lambda(K) - k(\lambda, \lambda)\} \qquad (\lambda \in \mathfrak{M}_K^+), \text{ cf. (1), p. 162},$$

are identical provided the pseudo-positive kernel k is *symmetric* and fulfills Frostman's *maximum principle*. In the special case of the Newtonian kernel, the common value is the classical capacity of K as defined first by N. Wiener [33] (for arbitrary compact sets). As to the history of the mathematical concept of (interior) Newtonian capacity, see O. Frostman [18], Chap. III. Some historical remarks concerning the *exterior* Newtonian capacity are found in M. Brelot [7], p. 136.

denote the interior and exterior Wiener capacity by cap_* and cap^*, respectively. The following characterization of $\text{cap}_* A$ is useful:

$$\text{cap}_* A = 1/w(A) = \sup \{2\lambda(X) - k(\lambda, \lambda)\}, \tag{1}$$

where λ ranges over the class of all positive measures of finite energy concentrated on A (or equivalently: supported by some compact subset of A). [1] As usual we may write cap A instead of $\text{cap}_* A$ or $\text{cap}^* A$ if A is capacitable.

From Theorems 2.3 and 2.4 we derive the following result, using the correspondence $\lambda = t \cdot \mu$ $(\mu(X) = 1)$ between the two classes of competing measures.

THEOREM 2.5. *Let k denote a symmetric, pseudo-positive kernel, and K a compact set with* cap $K < +\infty$. [2] *The two maximum problems*

$$\lambda(K) = \text{maximum} \quad (\lambda \in \mathfrak{M}_K^+, \; V(\lambda) \leqslant 1),$$

and

$$2\lambda(K) - k(\lambda, \lambda) = \text{maximum} \quad (\lambda \in \mathfrak{M}_K^+)$$

have precisely the same solutions, and the value of each of the two maxima is the Wiener capacity cap K. *The class of all solutions is compact in the vague topology on $\mathfrak{M}^+$ and consists of all measures $\lambda \in \mathfrak{M}_K^+$ for which*

$$k(\lambda, \lambda) = \lambda(X) = \text{cap}\, K.$$

The potential of any solution has the following properties:

(a) $k(x, \lambda) \geqslant 1$ *nearly everywhere in K.*

(b) $k(x, \lambda) \leqslant 1$ *everywhere in the support of λ.*

(c) $k(x, \lambda) = 1$ *λ-almost everywhere in X.*

The solutions mentioned above are called *capacitary distributions* on K. The class of all capacitary distributions on K is denoted by Λ_K' (cf. Theorem 4.1 and the remark following it). As to the proof of Theorem 2.5, the only point in need of comment is the fact that $\lambda \in \Lambda_K'$ if $\lambda \in \mathfrak{M}_K^+$ and $k(\lambda, \lambda) = \lambda(X) = \text{cap}\, K$; and this follows

---

(1) This formula follows from the fact that $\lambda(X)$ and $k(\lambda, \lambda)$ are homogeneous in λ of order 1 and 2, respectively. Disregarding the trivial case where $w(A) = +\infty$, we may restrict the attention to non-zero measures λ in (1), as it will appear presently. Writing $\lambda = t \cdot \mu$, where $\mu(X) = 1$ and $t = \lambda(X) > 0$, we obtain the quadratic $2t - k(\mu, \mu)t^2$, which attains its maximum at $t = k(\mu, \mu)^{-1}$, the maximum being $t = k(\mu, \mu)^{-1} > 0$. Maximizing over μ, we obtain $1/w(A)$.

(2) Recall that the condition cap $K < +\infty$ is fulfilled for *all* compact sets K if and only if the kernel k is *strictly pseudo-positive*; thus in particular if k is strictly positive ($k(x, y) \geqslant 0$ and $k(x, x) > 0$) or strictly definite.

at once from the implied relation $2\lambda(X)-k(\lambda,\lambda)=\text{cap } K$. Observe that Λ'_K reduces to the single measure $\lambda=0$ if and only if $\text{cap } K=0$.

The remarks following Theorem 2.4 can be carried over to the present situation. 1) If k fulfills Frostman's *maximum principle*, then $k(x,\lambda)=1$ n.e. in K, and $k(x,\lambda)\leqslant 1$ everywhere. 2) If k is *definite*, Λ'_K is convex and consists of all measures $\lambda\in\mathcal{E}^+_K$ whose potentials have the properties (a) and (b) (or equivalently (a) and (c)). Moreover, $k(\lambda-\mu,\lambda-\mu)=0$ for any two capacitary distributions λ and μ on K. If k is *strictly definite*, there is just one capacitary distribution on K.

II. The Case of a Consistent Kernel

In the present chapter we study potentials with respect to a *definite* kernel k on a locally compact space X. The limitation to the two cases $\text{I}: k\geqslant 0$, and $\text{II}: X$ compact (cf. § 2.1) remains in force in the present chapter, except for § 5.3.

Since a definite kernel is pseudo-positive, the interior Wiener capacity $\text{cap}_* A = 1/w(A) = 1/v(A)$ is defined for arbitrary sets $A\subset X$, cf. (1), § 2.5, and (7), § 2.3; and so is the exterior Wiener capacity $\text{cap}^* A = 1/w^*(A) = 1/v^*(A)$. According to (2) and (3), § 2.3,

$$\text{cap}_* A = \sup_{K\subset A} \text{cap } K; \qquad \text{cap}^* A = \inf_{G\supset A} \text{cap } G,$$

where K and G refer to compact and open sets, respectively. The "small" sets N in terms of which the concepts "nearly everywhere" (n.e.) and "quasi-everywhere" (q.-e.) were defined in § 2.3, are those for which $\text{cap}_* N=0$ and $\text{cap}^* N=0$, respectively. For any set $A\subset X$,

$$0\leqslant \text{cap}_* A\leqslant \text{cap}^* A\leqslant +\infty.$$

If $\text{cap}_* A=\text{cap}^* A$, we call A capacitable and may write simply cap A for the Wiener capacity of A.

The principal aim of the present chapter is to show (in § 4) that the concept and the properties of the capacitary distributions on a compact set (Theorem 2.5) can be extended in a satisfactory way to arbitrary sets of finite interior or exterior capacity, provided the definite kernel k is *consistent* (§ 3.3).[1] Except for the case of closed sets, one must, however, give up the requirement that the capacitary distributions should be concentrated on the set in question. The results will allow us to apply Choquet's theory of capacitability.

[1] The fact that some restriction on the definite kernel is indispensable appears from examples 4 and 5, § 8.3.

3. The strong topology

3.1. *The spaces $\mathcal{E}$ and $\mathcal{E}^+$.* Consider a definite kernel k (cf. § 2.1), and denote by $\mathcal{E} = \mathcal{E}(X)$ the class of all measures μ on X such that the energy $k(\mu, \mu)$ is defined and *finite* (i.e., $\neq +\infty$). The class of all positive measures $\mu \in \mathcal{E}$ is denoted by $\mathcal{E}^+ = \mathcal{E}^+(X)$.

LEMMA 3.1.1. *$\mathcal{E}^+$ is a convex cone. The mutual energy $k(\mu, \nu)$ is defined and finite when $\mu, \nu \in \mathcal{E}^+$.*

Proof. The former statement is easily derived from the latter, which in turn is implied by the fact that

$$k(\mu - \nu, \mu - \nu) = k(\mu, \mu) + k(\nu, \nu) - 2k(\mu, \nu)$$

is defined (because $k(\mu, \mu) + k(\nu, \nu) < +\infty$, cf. Lemma 1.2.3) and hence $\geqslant 0$.

If $\mu \in \mathcal{E}$, then $\mu^+, \mu^- \in \mathcal{E}^+$. The converse statement follows from the above lemma. In particular, $\mathcal{E} = \mathcal{E}^+ - \mathcal{E}^+$. Another consequence is that the mutual energy $k(\mu, \nu)$ is defined and finite for any two measures $\mu, \nu \in \mathcal{E}$, cf. (1), § 2.1. Moreover, $k(\mu, \nu)$ is a bilinear form on $\mathcal{E}$; and since $k(\mu, \mu) \geqslant 0$ for every $\mu \in \mathcal{E}$, we have obtained the following result:

LEMMA 3.1.2. *$\mathcal{E}$ is a pre-Hilbert space* (over the field of real numbers) *with the scalar product $k(\mu, \nu)$ and the semi-norm*

$$\|\mu\| = (k(\mu, \mu))^{\frac{1}{2}}.$$

This semi-norm is a *norm* (i.e., $\mathcal{E}$ is a Hausdorff space) if and only if the kernel k is *strictly definite.* As a corollary of Lemma 3.1.2 we obtain the Cauchy-Schwarz inequality:

$$|k(\mu, \nu)| \leqslant \|\mu\| \cdot \|\nu\| \qquad (\mu, \nu \in \mathcal{E}), \tag{1}$$

valid even if k is not strictly definite (cf., e.g., Bourbaki [6], Chap. V, § 1, prop. 2). Two measures λ and μ in $\mathcal{E}$ are called *equivalent* if $\|\lambda - \mu\| = 0$.

In addition to the *strong topology* on $\mathcal{E}$, defined by the above semi-norm $\|\mu\|$, it is sometimes useful to consider the *weak topology* on $\mathcal{E}$, defined by the semi-norms $\mu \to |k(\mu, \nu)|$, $\nu \in \mathcal{E}$. The induced topologies on $\mathcal{E}^+$ are likewise called strong and weak topologies on $\mathcal{E}^+$.

3.2. *Potentials with respect to a definite kernel.* If $\mu \in \mathcal{E}$, $k(x, \mu) \in \mathfrak{L}^1(\nu)$ for every $\nu \in \mathcal{E}$. Hence $k(x, \mu)$ is defined and finite n.e. in X (cf. Lemma 2.3.1).

LEMMA 3.2.1. *Each of the following two conditions is necessary and sufficient in order that two measures $\lambda, \mu \in \mathcal{E}$ be equivalent:*

(a) $k(x, \lambda) = k(x, \mu)$ for nearly every $x \in X$.

(a′) $k(\nu, \lambda) = k(\nu, \mu)$ for every $\nu \in \mathcal{E}$.

Proof. That (a′) implies $\|\lambda - \mu\| = 0$ follows when we take $\nu = \lambda - \mu$. The converse statement follows from the Cauchy-Schwarz inequality (1), § 3.1. The equivalence between (a) and (a′) is verified as in the proof of Lemma 2.3.2

COROLLARY. Any two capacitary distributions on a compact set (of finite capacity) have nearly everywhere the same potential. (In fact, the two capacitary distributions are equivalent in view of the second remark to Theorem 2.5.)

LEMMA 3.2.2. *Let $0 < t \leqslant +\infty$, $\mu \in \mathcal{E}$, and let A denote a set such that $k(x, \mu) \geqslant t$ nearly everywhere in A. Then*

$$\mathrm{cap}_* A \leqslant t^{-2} \|\mu\|^2.$$

Proof. Let λ denote a capacitary distribution on some compact set $K \subset A$. According to Lemma 2.3.2, which holds likewise for $\mu \in \mathcal{E}$,

$$t \cdot \lambda(X) \leqslant k(\lambda, \mu) \leqslant \|\lambda\| \, \|\mu\|.$$

Inserting $\lambda(X) = \|\lambda\|^2 = \mathrm{cap}\, K$ (cf. Theorem 2.5), we obtain $\mathrm{cap}\, K \leqslant t^{-2} \|\mu\|^2$, from which the stated inequality follows.

A similar lemma subsists for the *exterior* capacity, at least under certain further restrictions (cf. Lemma 4.3.2 and the note attached to it). In the case of a *positive* measure μ, no further restrictions are needed:

LEMMA 3.2.3. *Let $0 < t \leqslant +\infty$, $\mu \in \mathcal{E}^+$, and let A denote a set such that $k(x, \mu) \geqslant t$ quasi-everywhere in A. Then*

$$\mathrm{cap}^* A \leqslant t^{-2} \|\mu\|^2.$$

Proof. In view of the corollary to Lemma 2.3.5, we may assume without loss of generality that $k(x, \mu) \geqslant t$ *everywhere* in A. For any number s, $0 < s < t$, A is then contained in the *open* set $\{x \in X\colon k(x, \mu) > s\}$, whose capacity is $\leqslant s^{-2} \|\mu\|^2$ according to the preceding lemma. Hence $\mathrm{cap}^* A \leqslant s^{-2} \|\mu\|^2$, and the result is obtained by letting $s \to t$. Applying this lemma with $t = +\infty$ to μ^+ and μ^- $(\mu \in \mathcal{E})$, we obtain:

COROLLARY. The potential $k(x, \mu)$ of any measure $\mu \in \mathcal{E}$ is defined and finite quasi-everywhere in X.

LEMMA 3.2.4. *If $\mu_n \to \mu$ strongly in $\mathcal{E}$, then*

$$k(x, \mu) \geqslant \liminf_n k(x, \mu_n)$$

nearly everywhere in X.

Proof. Let X' denote the set of points x at which $k(x, \mu)$ and all $k(x, \mu_n)$ are defined and finite. The set of points at which the stated inequality does not hold is contained in $N \cup \complement X'$, where

$$N = \{x \in X' : k(x, \mu) < \liminf k(x, \mu_n)\}.$$

Writing
$$N_{p,q} = \{x \in X' : k(x, \mu) \leqslant \inf_{n>p} k(x, \mu_n) - 1/q\},$$

$$A_{n,q} = \{x \in X' : k(x, \mu) \leqslant k(x, \mu_n) - 1/q\},$$

we obtain
$$N = \bigcup_{p,q} N_{p,q}; \qquad N_{p,q} = \bigcap_{n>p} A_{n,q}.$$

According to the corollaries to Lemmas 2.3.5 and 3.2.3, it suffices to prove that $\text{cap}_* N_{p,q} = 0$; and this follows from Lemma 3.2.2 with t, μ, and A replaced by $1/q$, $\mu_n - \mu$, and $A_{n,q}$, respectively:

$$\text{cap}_* N_{p,q} \leqslant \text{cap}_* A_{n,q} \leqslant q^2 \|\mu_n - \mu\|^2 \to 0$$

as $n \to \infty$.

Remark. It is not difficult to prove the following stronger result: *If $\mu_n \to \mu$ strongly in $\mathcal{E}$, there is a subsequence $\{\mu_{n_i}\}$ whose potentials converge to the potential $k(x, \mu)$ of μ nearly everywhere.*

3.3. *Perfect kernel. Consistent kernel.* Simple examples show that the pre-Hilbert space $\mathcal{E}$ is in general incomplete (in the uniform structure defined by the "semi-distance" $\|\mu - \nu\|$). In case of the Newtonian kernel $|x-y|^{2-n}$ in R^n $(n \geqslant 3)$ it was proved by H. Cartan [10], § 5, that the space $\mathcal{E}^+$ of *positive* measures of finite energy is strongly complete, and that strong convergence in $\mathcal{E}^+$ implies vague convergence to the same limit.

DEFINITION. A definite kernel is called *perfect* if the following two conditions are fulfilled: (1)

(P_1) *$\mathcal{E}^+$ is strongly complete,* i.e., any strong Cauchy filter on $\mathcal{E}^+$ converges strongly in $\mathcal{E}^+$.

(P_2) *The strong topology on $\mathcal{E}^+$ is finer than the induced vague togology on $\mathcal{E}^+$,* i.e., any strongly convergent filter on $\mathcal{E}^+$ converges vaguely to the same limit.

(1) The conditions (P_1) and (P_2) do not follow from one another (cf. Ex. 1 and Ex. 4 or 5, § 8.3). It is not known to the author whether (P_2) follows from (P_1) in the case of a *strictly* definite kernel. Note that Ex. 5 shows that $\mathcal{E}^+$ may be incomplete in case of a positive and strictly definite kernel, even on a compact space. (This cannot occur in case of convolution kernels, cf. the corollary to Theorem 7.2). Finally we observe that (P_1), the strong completeness of $\mathcal{E}^+$, may be expressed as follows: Every strong Cauchy *sequence* in $\mathcal{E}^+$ converges strongly (cf. the end of the proof of Lemma 3.3.2).

A definite kernel possessing property (P_2) is necessarily *strictly definite.* (In fact, if $\mu \in \mathcal{E}$ and $\|\mu\| = 0$, then $\|\mu^+ - \mu^-\| = 0$, and the sequence $\mu^+, \mu^+, \ldots$ converges strongly, hence vaguely, to μ^-. Since the vague topology is separated, we conclude that $\mu^+ = \mu^-$.) For the subsequent applications the property of *strict* definiteness is of minor importance. We shall therefore introduce a concept similar to perfectness, but applicable even to definite kernels which are not strictly definite. The crucial property is a kind of consistency between the strong (not necessarily separated) topology and the vague topology on $\mathcal{E}^+$:

DEFINITION. A definite kernel is called *consistent* if the following condition is fulfilled:

(C) If μ is a vague cluster point for a strong Cauchy filter Φ on $\mathcal{E}^+$, then Φ converges strongly to μ.

The following apparently weaker condition is equivalent to (C):

(C′) If a strong Cauchy filter Φ on $\mathcal{E}^+$ converges vaguely to μ, then Φ converges strongly to μ.

In fact, if μ adheres vaguely to a strong Cauchy filter Φ, there is a finer filter $\Phi' \supset \Phi$ which converges vaguely to μ. According to (C′), $\Phi' \to \mu$ strongly, and hence $\Phi \to \mu$ strongly because Φ is a strong Cauchy filter (cf. Bourbaki [2], Chap. II, § 3, prop. 4).

It follows immediately from (C) that any strong Cauchy filter Φ possessing a basis consisting of vaguely bounded subsets of $\mathcal{E}^+$, converges strongly. (The vague adherence of Φ is, in fact, non-void because any vaguely bounded subset of $\mathfrak{M}^+$ is vaguely relatively compact.) If a consistent kernel is strictly pseudo-positive, every strongly bounded subset of $\mathcal{E}^+$ is vaguely bounded (Lemma 2.5.1), so that we obtain the following result:

LEMMA 3.3.1. *If the kernel is consistent and strictly pseudo-positive, the space $\mathcal{E}^+$ is strongly complete.*

The consistency alone is not sufficient for $\mathcal{E}^+$ to be complete (cf. Ex. 6, § 8.3). On the other hand, the example $k = 0$ shows that $\mathcal{E}^+$ may be complete in the case of a consistent kernel which is not strictly pseudo-positive.

THEOREM 3.3. *A kernel is perfect if and only if it is consistent and strictly definite.*

Proof. Suppose first k is consistent and strictly definite. Then k is strictly pseudo-positive, and hence $\mathcal{E}^+$ is complete according to the above lemma. If a filter Φ converges strongly, the strong limit μ is uniquely determined (because $\mathcal{E}^+$ is separated in the strong topology when k is strictly definite). The consistency of k implies that Φ can have no other vague cluster points than μ. Since Φ possesses a basis consisting of vaguely relatively compact sets, we conclude that the vague adherence of Φ is non-void and reduces to the single measure μ; and consequently $\Phi \to \mu$ vaguely (cf. Bourbaki [2], Chap. I, § 10, N° 1, cor.). The converse is obvious when we use (C′) as a definition of consistency: If k is perfect, and if a strong Cauchy filter Φ on $\mathcal{E}^+$ converges vaguely to μ, then (P_1) implies that Φ converges strongly, and (P_2) shows that the strong limit is μ.

LEMMA 3.3.2. *Suppose the locally compact space X is metrizable and of class K_σ. The concepts of a perfect or a consistent kernel remain unchanged if, in either of the definitions* (P_1), (P_2), *resp.* (C) *or* (C′), *the filter Φ is replaced by a sequence.*

Proof. We begin by considering property (C′). It is assumed that any vaguely convergent, strong Cauchy *sequence* on $\mathcal{E}^+$ converges strongly to its vague limit. We propose to show that the kernel is consistent according to definition (C). Thus we consider an arbitrary vague cluster point μ_0 for a strong Cauchy filter Φ on $\mathcal{E}^+$. Since a strong Cauchy filter contains sets of arbitrarily small diameters, there are sets $A_n \in \Phi$ such that $A_n \subset A_p$ for $n > p$, and diam $A_n \to 0$ as $n \to \infty$. According to Lemma 1.2.1 there is a sequence of sets $V_n \subset \mathfrak{M}^+$ forming a fundamental system of neighbourhoods of μ_0 in the vague topology on $\mathfrak{M}^+$. Each vague neighbourhood of μ_0 intersects every set $A \in \Phi$ because μ_0 adheres vaguely to Φ. Choose $\mu_n \in A_n \cap V_n$. For $n > p$ we have $\mu_n \in A_p$, and hence

$$\|\mu_n - \mu_p\| \leqslant \operatorname{diam} A_p \qquad (n > p).$$

Consequently, $\{\mu_n\}$ is a strong Cauchy sequence. Furthermore, $\mu_n \to \mu_0$ vaguely because $\mu_n \in V_n$. By hypothesis, $\mu_n \to \mu_0$ strongly. Since

$$\sup_{\mu \in A_n} \|\mu - \mu_0\| \leqslant \|\mu_n - \mu_0\| + \operatorname{diam} A_n,$$

we conclude that $\Phi \to \mu_0$ strongly. This completes the proof in case of condition (C) or (C′). Next, consider a definite kernel k such that (i) every strong Cauchy sequence in $\mathcal{E}^+$ converges strongly, and (ii) every strongly convergent sequence in $\mathcal{E}^+$ converges vaguely to the same limit. The simple argument at the end of the proof of Theorem 3.3 shows that such a kernel has the property corresponding to (C′), but with a sequence instead of a filter. We have, however, just shown that this property implies consistency, and since k is strictly definite (same argument as earlier), we conclude from Theorem 3.3 that k is perfect. Note also that the above reasoning with the sets A_n, etc., shows that $\mathcal{E}^+$ is complete if the kernel has the property that every strong Cauchy *sequence* in $\mathcal{E}^+$ converges strongly.

The following diagram illustrates the relations between various types of kernels introduced in the preceding sections (in the cases I : $k \geqslant 0$, or II : X compact).

$$\begin{array}{ccccccc}
\text{perfect} & \Rightarrow & \begin{array}{c}\text{strictly}\\ \text{definite}\end{array} & \Rightarrow & \begin{array}{c}\text{strictly}\\ \text{pseudo-positive}\end{array} & \Leftarrow & \begin{array}{c}\text{strictly}\\ \text{positive}\end{array} \\
\Downarrow & & \Downarrow & & \Downarrow & & \Downarrow \\
\text{consistent} & \Rightarrow & \text{definite} & \Rightarrow & \text{pseudo-positive} & \Leftarrow & \text{positive} \\
 & & \Downarrow & & & & \\
 & & \text{symmetric} & & & &
\end{array}$$

3.4. *Criteria for consistency or strict definiteness.* The following three lemmas can be extracted from Cartan's proof of the perfectness of the Newtonian kernel (H. Cartan [10], §§ 4, 5). We begin by discussing a kind of consistency between the vague and the *weak* topologies on $\mathcal{E}^+$.

LEMMA 3.4.1. *A definite kernel possessing the following property is consistent:*

(CW) If a filter Φ on a strongly bounded part of $\mathcal{E}^+$ converges vaguely to μ, then Φ converges weakly to μ.

Proof. It is well known that a Cauchy filter Φ on a pre-Hilbert space has a base formed by subsets of a strongly bounded set, and that, if Φ converges weakly, then Φ converges strongly to the same limit. (1)

LEMMA 3.4.2. *A sufficient condition for a definite kernel k to be consistent is that the class of all measures $\lambda \in \mathcal{E}$ for which the potential $k(x, \lambda)$ is of class $C_0(X)$ be strongly dense in $\mathcal{E}$.*

---

(1) Let Φ denote a strong Cauchy filter on a pre-Hilbert space $\mathcal{H}$ (say, over the field of real numbers) with elements x, y, z, etc.; scalar product (x, y); and semi-norm $\|x\| = (x, x)^{\frac{1}{2}}$. Suppose Φ converges weakly to some vector x_0, that is,

$$\lim_x \,(x, z) = (x_0, z) \text{ along } \Phi$$

for every $z \in \mathcal{H}$. Then $\|x_0\| \leqslant \liminf \|x\|$ along Φ because

$$\|x_0\|^2 = \lim_x \,(x, x_0) \leqslant \liminf_x \|x\| \cdot \|x_0\|.$$

Hence,

$$\|x - x_0\|^2 = \|x\|^2 + \|x_0\|^2 - 2\,(x, x_0) \leqslant \liminf_y \left\{\|x\|^2 + \|y\|^2 - 2\,(x, y)\right\} = \liminf_y \|x - y\|^2.$$

The function of x on the right approaches 0 along Φ.

Proof. We propose to verify condition (CW). Let Φ denote a filter on the part of $\mathcal{E}^+$ determined by the inequality $\|\mu\|^2 \leqslant M$, and suppose $\Phi \to \mu_0$ vaguely. Then

$$k(\mu, \lambda) = \int k(x, \lambda)\, d\mu \to \int k(x, \lambda)\, d\mu_0 = k(\mu_0, \lambda)$$

for every measure $\lambda \in \mathcal{E}$ such that $k(x, \lambda) \in \mathcal{C}_0$. In view of the hypothesis and the boundedness condition $\|\mu\|^2 \leqslant M$, this implies that $k(\mu, \lambda) \to k(\mu_0, \lambda)$ for *every* $\lambda \in \mathcal{E}$, i.e., $\Phi \to \mu_0$ weakly.

LEMMA 3.4.3. *A sufficient condition for a definite kernel k to be strictly definite is that the class of all functions $f \in \mathcal{C}_0(X)$ representable as potentials $k(x, \lambda)$ with $\lambda \in \mathcal{E}$ be rich in $\mathcal{C}_0(X)$.*

(As to the notion of a *rich* subclass of $\mathcal{C}_0$, see Bourbaki [4], Chap. III, § 2, N° 5.) In fact, if $\|\mu\| = 0$ for some $\mu \in \mathcal{E}$, it follows from the Cauchy-Schwarz inequality that $\int k(x, \lambda)\, d\mu = k(\mu, \lambda) = 0$ for every $\lambda \in \mathcal{E}$. Hence $\mu(f) = \int f\, d\mu = 0$ for every f of the type described in the lemma, and it follows that $\mu = 0$ (cf. Bourbaki [4], loc. cit.).

Remark. The criteria for consistency or strict definiteness formulated in Lemmas 3.4.2 and 3.4.3 are not necessary conditions (except possibly in case II where X is compact). This appears from Exs. 8 and 9, § 8.3. They are, however, fulfilled (and rather easily verified) by many interesting kernels, e.g. Green's function for the Laplace operator, in particular the Newtonian kernel; furthermore, the kernels of order α, $0 < \alpha < n$; and also the kernels considered by J. Deny [15] (inasmuch as the definition of energy given there agrees with the classical definition used in the present paper, cf. Deny [16] and also Theorem 7.3 in the present paper). Finally, it was shown by M. Ohtsuka [24] that the criterion given in Lemma 3.4.2 is fulfilled by any regular, definite kernel on a compact space: (1)

THEOREM 3.4.1. *Every regular, definite kernel k on a compact space X is consistent. To every measure $\mu \in \mathcal{E}^+$ and every number $\varepsilon > 0$ corresponds a measure $\lambda \in \mathcal{E}^+$ such that $\lambda \leqslant \mu$, $k(x, \lambda)$ is continuous in X, and $\|\lambda - \mu\| < \varepsilon$.*

Proof. Let c denote a constant $\geqslant 0$ such that $k + c \geqslant 0$, and write

$$f(x) = \int (k(x, y) + c)\, d\mu(y) = k(x, \mu) + c \cdot \mu(X).$$

---

(1) The theorem of Ohtsuka states that $\mathcal{E}^+$ is complete if k is positive, regular, and *strictly* definite (and if X is compact); but his proof shows that such a kernel is consistent (and hence perfect according to Theorem 3.3). As to the second assertion of the theorem, see also H. Cartan [10], lemme 5, p. 98, in the Newtonian case; and G. Choquet [13] in the general case.

Since $k(x, \mu)$ is μ-integrable, so is f, and hence we may introduce the measure $\nu = f \cdot \mu$ of density f with respect to μ. For the μ-measurable set $N = \{x \in X: k(x, \mu) = +\infty\}$, we have $\mu(N) = 0$ because $k(x, \mu)$ is μ-integrable. Hence $\nu(N) = \int_N f\,d\mu = 0$. According to (1), § 1.1, there is a compact set $K \subset \complement N$ such that

$$\nu(K) > \nu(\complement N) - \varepsilon = \nu(X) - \varepsilon,$$

the number $\varepsilon > 0$ being given. By virtue of Lusin's theorem, [1] there is a compact set $K_1 \subset K$ such that the restriction of $k(x, \mu)$ to K_1 is continuous (in the usual sense because $k(x, \mu)$ is *finite* on K), and

$$\nu(K \cap \complement K_1) < \varepsilon.$$

It follows that
$$\int_{\complement K_1} f\,d\mu = \nu(\complement K_1) < 2\,\varepsilon.$$

Let λ denote the trace of μ on K_1. Clearly, $\lambda \in \mathcal{E}^+$, $S(\lambda) \subset K_1$, and the restriction of $k(x, \lambda)$ to K_1 is likewise continuous. (It is evidently finite and lower semi-continuous, and the equation

$$k(x, \lambda) = k(x, \mu) - k(x, \mu - \lambda)$$

shows that it is upper semi-continuous because $\mu - \lambda \geqslant 0$.) The regularity of k implies now that $k(x, \lambda)$ is continuous throughout X. Finally,

$$k(x, \mu - \lambda) = k(x, \mu) - k(x, \lambda) \leqslant k(x, \mu) + c \cdot \lambda(X) \leqslant f,$$

and hence
$$\|\mu - \lambda\|^2 = \int_{\complement K_1} k(x, \mu - \lambda)\,d\mu(x) \leqslant \int_{\complement K_1} f\,d\mu < 2\,\varepsilon.$$

In the following theorem the concept of a kernel k on a space X is taken in the restricted sense in which it occurs e.g. in the Japanese literature. It is required that k is a *continuous* mapping of $X \times X$ into the *extended* real line; that $k(x, y) \neq +\infty$ for $x \neq y$, and finally that $k(x, y) \neq -\infty$. We shall, moreover, call a kernel K-consistent if its restriction to $K \times K$ is consistent for every compact set $K \subset X$ (cf. § 5.3).

THEOREM 3.4.2. *On a locally compact space X, any symmetric kernel $k \geqslant 0$ satisfying Frostman's maximum principle is K-consistent.*

Proof. The kernel k is definite according to an interesting result due to N. Ninomiya [23], th. 3; and k is regular as shown by G. Choquet [12] and N. Nino-

[1] The idea of applying Lusin's or Egorov's theorem in potential theory seems to go back to K. Yosida [34].

miya [23], Lemme 3. Hence we obtain the K-consistency of k by application of Theorem 3.4.1 to the restriction of k to $K \times K$ for arbitrary compact $K \subset X$. Observe also that the space $\mathcal{E}_K^+$ of all positive measures of finite energy with respect to k, and supported by K, is strongly *complete* for every compact $K \subset X$ if we assume, in addition, that k is *strictly* positive: $k(x, x) > 0$ for every $x \in X$.

3.5. *Superposition of kernels.* The totality of all consistent kernels k on X (where $k \geqslant 0$ unless X is compact) is easily seen to form a convex cone. Even an infinite sum of consistent kernels $k \geqslant 0$ is consistent, as one may show by the method used in the proof of Theorem 3.5 below. We shall now consider a more general type of superposition of kernels. Let X and T denote two locally compact Hausdorff spaces, and let $\tau \geqslant 0$ be a fixed measure on T. Further, let $k(x, y, t)$ denote a lower semicontinuous function of (x, y, t) defined on $X \times X \times T$, and suppose $k(x, y, t) \geqslant 0$ everywhere unless X and T are compact. The function k on $X \times X$ defined by

$$k(x, y) = \int_T k(x, y, t)\, d\tau(t)$$

is then a kernel on X obtained, as we say, by *superposition* of the family of kernels $k_t = k(x, y, t)$, $t \in T$. Consider now any two measures μ and ν on X whose mutual energy $k(\mu, \nu)$ with respect to the kernel $k = \int k_t\, d\tau$ is defined. With the notation $\lambda = \mu \otimes \nu$, this means that $\int k\, d\lambda^+$ and $\int k\, d\lambda^-$ are not both infinite. Applying case (ii) of Fubini's theorem (§ 1.1) to the positive measures λ^+, resp. λ^-, and τ, we obtain

$$\int k\, d\lambda^+ = \int d\tau(t) \int k_t\, d\lambda^+,$$

$$\int k\, d\lambda^- = \int d\tau(t) \int k_t\, d\lambda^-.$$

Thus one, at least, of the functions $\int k_t\, d\lambda^+$ and $\int k_t\, d\lambda^-$ is τ-integrable and hence finite τ-almost everywhere. This shows that the mutual energy $k_t(\mu, \nu) = \int k_t\, d\lambda$ of μ and ν with respect to the kernel k_t is defined for τ-a.e. $t \in T$; and also that $\int k\, d\lambda = \int d\tau(t) \int k_t\, d\lambda$, i.e.,

$$k(\mu, \nu) = \int k_t(\mu, \nu)\, d\tau(t). \tag{1}$$

We shall now assume that k_t is *definite* for τ-a.e. $t \in T$. In view of (1) this implies that k is definite. Moreover, any measure μ of finite energy $k(\mu, \mu) = \|\mu\|^2$ with respect to k is likewise of finite energy $k_t(\mu, \mu) = \|\mu\|_t^2$ with respect to k_t for τ-a.e. $t \in T$. Note also that k is *strictly definite* if $\tau^*(S) > 0$, where S denotes the set of values of t for which k_t is strictly definite. Except in the case where T is discrete (and hence k a *sum* of kernels k_t) it is not known to the author whether in general k is *consistent* if k_t is consistent for τ-a.e. $t \in T$. The following theorem answers this question in the affirmative provided X is metrizable and of class K_σ.

THEOREM 3.5. *On a locally compact, metrizable space X of class K_σ any kernel obtained in the above manner by superposition of consistent kernels is consistent.*

Proof. Consider a strong Cauchy sequence in $\mathcal{E}^+$ (where $\mathcal{E}^+$ is formed with respect to the given kernel k), and suppose this sequence converges vaguely to some measure μ. Our task is to prove that the given Cauchy sequence converges strongly to μ (cf. Lemma 3.3.2). We begin by choosing a subsequence $\{\mu_n\}$ of the given sequence so that $\|\mu_n - \mu_{n+1}\| \leq 2^{-n}$; that is, in view of (1),

$$\|\mu_n - \mu_{n+1}\|^2 = \int \|\mu_n - \mu_{n+1}\|_t^2 \, d\tau(t) \leq 4^{-n}.$$

Denoting by A_n the set of those $t \in T$ for which $\|\mu_n - \mu_{n+1}\|_t^2 \geq 2^{-n}$, we obtain $\tau(A_n) \leq 2^{-n}$. Writing $N = \bigcap_p \bigcup_{n>p} A_n$, we infer that $\tau(N) = 0$. This implies that $\{\mu_n\}$ is a strong Cauchy sequence in $\mathcal{E}_t^+$ (with respect to k_t) for τ-a.e. $t \in T$. Since $\mu_n \to \mu$ vaguely, and k_t is consistent, it follows that $\mu_n \to \mu$ strongly in $\mathcal{E}_t^+$ for τ-a.e. $t \in T$, and hence we have for every p

$$\lim_n \|\mu_p - \mu_n\|_t^2 = \|\mu_p - \mu\|_t^2 \qquad \text{for } \tau\text{-a.e. } t \in T.$$

An application of Fatou's theorem (Bourbaki [4], Chap. IV, § 1, prop. 14) gives

$$\|\mu_p - \mu\|^2 = \int \lim_n \|\mu_p - \mu_n\|_t^2 \, d\tau(t) \leq \liminf_n \|\mu_p - \mu_n\|^2.$$

Since $\{\mu_n\}$ is a Cauchy sequence (with respect to k), the expression on the right approaches 0 as $p \to \infty$. Consequently $\{\mu_n\}$, and hence also the given Cauchy sequence, converges strongly to μ.

4. The interior and exterior capacitary distributions

4.1. *The interior capacitary distributions.* We shall use the following elementary lemma from the geometry of pre-Hilbert spaces. Let Γ denote a convex subset of a pre-Hilbert space $\mathcal{H}$ with the scalar product (μ, ν) and the seminorm $\|\mu\|$. Consider the quantity

$$\|\Gamma\| = \inf_{\mu \in \Gamma} \|\mu\|,$$

interpreted as $+\infty$ if Γ is void. With these notations one has the following lemma.

LEMMA 4.1.1. *If Γ contains some vector λ of minimal norm: $\|\lambda\| = \|\Gamma\|$, then the totality of all such minimal vectors λ is an equivalence class in $\mathcal{H}$. The inequality*

$$\|\mu - \lambda\|^2 \leq \|\mu\|^2 - \|\lambda\|^2 \tag{1}$$

holds for any $\mu \in \Gamma$ and any minimal $\lambda \in \Gamma$.

Proof. It suffices to establish (1). For any number t, $0 \leq t \leq 1$, the vector

$$\nu = (1-t) \cdot \lambda + t \cdot \mu = \lambda + t(\mu - \lambda)$$

belongs to Γ, and hence $\|\nu\|^2 \geq \|\lambda\|^2$. Evaluating $\|\nu\|^2$, we obtain $(\mu - \lambda, \lambda) \geq 0$, and (1) follows.

We shall apply this lemma to the pre-Hilbert space $\mathcal{E}$ consisting of all measures μ of finite energy $k(\mu, \mu) = \|\mu\|^2$ with respect to a given definite kernel k on a locally compact space X. For any set $A \subset X$ we denote by Γ_A the convex class of all measures $\mu \in \mathcal{E}$ with the property

(a) $k(x, \mu) \geq 1$ nearly everywhere in A,

or, equivalently (cf. Lemma 2.3.2, which holds likewise for $\mu \in \mathcal{E}$),

(a′) $k(\nu, \mu) \geq \nu(X)$ for every $\nu \in \mathcal{E}^+$ supported by some compact subset of A.

In view of, say, this latter definition, the convex set Γ_A is strongly closed in $\mathcal{E}$. According to Lemma 3.2.2 (with $t = 1$), $\|\mu\|^2 \geq \mathrm{cap}_* A$ for every $\mu \in \Gamma_A$, that is,

$$\|\Gamma_A\|^2 \geq \mathrm{cap}_* A. \tag{2}$$

In particular, Γ_A is void unless $\mathrm{cap}_* A < +\infty$. We proceed to prove that, actually,

$$\|\Gamma_A\|^2 = \mathrm{cap}_* A \tag{3}$$

provided the kernel k is *consistent*. In view of (2), we may suppose $\mathrm{cap}_* A < +\infty$, and it suffices to establish the existence of a single measure $\lambda \in \Gamma_A$ with $\|\lambda\|^2 = \mathrm{cap}_* A$. Hence (3) is contained in the following theorem.

THEOREM 4.1. *Suppose the kernel k is consistent, and let $A \subset X$, $\mathrm{cap}_* A < +\infty$ The class Λ_A of all positive measures λ for which*

$$\|\lambda\|^2 = \lambda(X) = \mathrm{cap}_* A,$$

and (a) $k(x, \lambda) \geqslant 1$ *nearly everywhere in* A,

(b) $k(x, \lambda) \leqslant 1$ *everywhere in the support of* λ,

and hence (c) $k(x, \lambda) = 1$ λ*-almost everywhere in* X,

is non-void and vaguely compact.

Remarks. The same is true of the class Λ'_A of all measures $\lambda \in \Lambda_A$ supported by the closure $\bar{A}$ of A. The measures forming the class Λ_A are called the *interior capacitary distributions* associated with the set A. In general, none of them is concentrated on A (unless A is closed). (1) For a *closed* set F with $\text{cap}_* F < +\infty$, the measures $\lambda \in \Lambda'_F$ may be called interior capacitary distributions *on* F. They have all the properties listed in Theorem 2.5 for the capacitary distributions on a compact set (provided capacity is interpreted as interior capacity), and coincide with these if F is compact. According to Lemma 4.1.1, Λ_A (and hence Λ'_A) is contained in the *equivalence class* of all minimal measures in the convex class Γ_A. If k is *strictly definite* (and hence *perfect*, cf. Theorem 3.3), there is just one interior capacitary distribution associated with A, and this unique minimal measure in Γ_A is supported by $\bar{A}$. Finally we note that, in case of a consistent kernel satisfying Frostman's *maximum principle*, properties (a) and (b) above may be replaced by the following properties: $k(x, \lambda) = 1$ nearly everywhere in A, and $k(x, \lambda) \leqslant 1$ everywhere in X.

Proof of Theorem 4.1. The existence of a measure with the desired properties is proved by an approximation of A by means of compact subsets in the manner indicated by H. Cartan [10], p. 94 f., for the Newtonian kernel (cf. note (1), p. 176). We employ here a more elementary version of this construction (cf. De la Vallée-Poussin [30], Chap. II, § 6). Choose a sequence of compact subsets $K_n \subset A$ so that

$$\lim_n \text{cap } K_n = \text{cap}_* A. \tag{4}$$

Replacing, if necessary, K_n by $K_1 \cup K_2 \cup \cdots \cup K_n$, we may suppose $K_n \supset K_p$ for $n > p$. Denote by λ_n a capacitary distribution on K_n, and observe that

$$\lambda_n \in \Gamma_{K_p} \qquad \text{for } n > p. \tag{5}$$

It follows from Theorem 2.5 that λ_p is a minimal measure in Γ_{K_p}, and hence Lemma 4.1.1 implies

(1) In case of the Newtonian kernel $|x-y|^{-1}$ on R^3, the (unique) interior capacitary distribution λ associated with the open unit ball B coincides with the capacitary distribution on the closure $\bar{B}$ of B. Thus λ is the uniform distribution of unit mass on the unit *sphere*, and $\lambda(B) = 0$.

$$\|\lambda_n - \lambda_p\|^2 \leqslant \|\lambda_n\|^2 - \|\lambda_p\|^2 = \text{cap } K_n - \text{cap } K_p.$$

Since $\text{cap}_* A < +\infty$, we conclude in view of (4) that $\{\lambda_n\}$ is a strong Cauchy sequence in $\mathcal{E}^+$. As $\lambda_n(X) = \text{cap } K_n$ remains bounded, there exists a vague cluster point λ for the sequence $\{\lambda_n\}$. The consistency of k implies now that $\lambda_n \to \lambda$ strongly. Since Γ_{K_p} is strongly closed, we infer from (5) that $\lambda \in \Gamma_{K_p}$ for every p. In other words, $k(x, \lambda) \geqslant 1$ nearly everywhere in each K_p, and hence in the union

$$H = \bigcup_n K_n;$$

that is, $\lambda \in \Gamma_H$. Furthermore,

$$\|\lambda\|^2 = \lim_n \|\lambda_n\|^2 = \lim_n \text{cap } K_n \leqslant \text{cap}_* H \leqslant \text{cap}_* A.$$

Here the sign of equality subsists in view of (4):

$$\|\lambda\|^2 = \text{cap}_* H = \text{cap}_* A. \tag{6}$$

The fact that λ adheres *vaguely* to the sequence $\{\lambda_n\}$ implies $\lambda \geqslant 0$, $S(\lambda) \subset \bar{H}$ (because $S(\lambda_n) \subset K_n \subset \bar{H}$), $V(\lambda) \leqslant 1$ (in view of the lower semi-continuity of $V(\mu)$ on $\mathfrak{M}^+$, cf. Lemma 2.2.1), and similarly

$$\lambda(X) \leqslant \limsup_n \lambda_n(X) = \lim_n \text{cap } K_n = \text{cap}_* A. \tag{7}$$

Since $k(x, \lambda) \leqslant 1$ in $S(\lambda)$, we have $k(x, \lambda) \leqslant 1$ λ-almost everywhere, and hence

$$\text{cap}_* A = \|\lambda\|^2 = \int k(x, \lambda)\, d\lambda \leqslant \lambda(X). \tag{8}$$

Combining (7) and (8), we infer that, actually, $\lambda(X) = \text{cap}_* A$, and $k(x, \lambda) = 1$ λ-almost everywhere. All the properties of λ stated in the theorem have thus been established except that property (a) has been proved only with H in place of A. In order to prove that $k(x, \lambda) \geqslant 1$ nearly everywhere in A, we must show that cap $K = 0$ if K is compact, $K \subset A$, and $k(x, \lambda) < 1$ everywhere in K. Writing [1]

$$K'_n = K_n \cup K; \qquad H' = H \cup K = \bigcup_n K'_n,$$

we obtain $H \subset H' \subset A$, and hence, by (6), $\text{cap}_* H = \text{cap}_* H' = \text{cap}_* A$. If λ' is determined from the sequence $\{K'_n\}$ in the same way as λ was determined from the se-

(1) This extra consideration, serving to extend property (a) from H to A, is, of course, unnecessary if A is of class K_σ. However, it may be avoided even by an arbitrary set A if one operates with the filtering *family* of all compact subsets $K \subset A$ instead of the random sequence K_n.

quence $\{K_n\}$, we have $\|\lambda'\|^2 = \text{cap}_* H' = \text{cap}_* H$, and $\lambda' \in \Gamma_{H'} \subset \Gamma_H$ because $H' \supset H$. Hence λ' is minimal in Γ_H, and it follows from Lemma 4.1.1 that $\|\lambda' - \lambda\| = 0$, or equivalently (cf. Lemma 3.2.1)

$$k(x, \lambda) = k(x, \lambda') \text{ nearly everywhere in } X.$$

Since $K \subset H'$, we have $k(x, \lambda') \geqslant 1$ nearly everywhere in K, and hence $k(x, \lambda) \geqslant 1$ n.e. in K. By definition of K, $k(x, \lambda) < 1$ everywhere in K, and consequently $\text{cap}\, K = 0$.

Having thus proved that Λ'_A (and hence Λ_A) is non-void, we proceed to show that the classes Λ_A and Λ'_A, which are obviously vaguely relatively compact, are *closed* in the vague topology. If a measure μ adheres vaguely to Λ_A, there is a filter Φ on $\mathcal{E}^+$ converging vaguely to μ and possessing a base formed by subsets of Λ_A. Since Λ_A is contained in an equivalence class, Φ is a strong Cauchy filter. In view of the consistency of k, Φ converges strongly to μ, and hence μ is a minimal measure in Γ_A. The vague convergence of Φ to μ implies, just as above, that $\mu \geqslant 0$, $V(\mu) \leqslant 1$, $\mu(X) \leqslant \text{cap}_* A$, and, subsequently, $\mu(X) = \text{cap}_* A$. Consequently, μ is an interior capacitary distribution associated with A, and so Λ_A is vaguely closed. This implies that Λ'_A is vaguely closed, and the proof of Theorem 4.1 is complete.

LEMMA 4.1.2. *Every set of class K_σ is capacitable.*

It suffices to consider a set A with $\text{cap}_* A < +\infty$. Each interior capacitary distribution λ associated with A has a potential which is $\geqslant 1$ in A except in the set

$$N = A \cap \bigcup_n \{x \in X:\ k(x, \lambda) \leqslant 1 - n^{-1}\},$$

for which $\text{cap}_* N = 0$. When A is of class K_σ, so is N; say $N = \bigcup_p K_p$, where each K_p is compact, and $\text{cap}_* K_p = 0$. Since compact sets are capacitable (Lemma 2.3.4), we infer from the corollary to Lemma 2.3.5 that $\text{cap}^* N = 0$, that is, $k(x, \lambda) \geqslant 1$ quasi-everywhere in A. In view of Lemma 3.2.3 with $t = 1$, we conclude that

$$\text{cap}^* A \leqslant \|\lambda\|^2 = \text{cap}_* A.$$

4.2. *Monotone families of sets, and the associated interior capacitary distributions.* The first part of the following theorem was established earlier in the case of an arbitrary kernel (Lemma 2.3.3), but no use will be made of this fact. Recall that $\mathfrak{A}$ denotes the system of all sets $A \subset X$ which are measurable with respect to every measure on X.

THEOREM 4.2. *Suppose the kernel k is consistent. If A denotes the union of an increasing sequence of sets A_n of class $\mathfrak{A}$, then*

$$\text{cap}_* A = \lim_n \text{cap}_* A_n.$$

If, in addition, $\text{cap}_* A < +\infty$, *and if* λ_n *denotes an interior capacitary distribution associated with* A_n, *then every vague cluster point* λ *of the vaguely bounded sequence* $\{\lambda_n\}$ *is an interior capacitary distribution associated with* A, *and* $\lambda_n \to \lambda$ *strongly.* (1)

It follows that the equivalence class of all minimal measures in Γ_{A_n} converges strongly to the equivalence class of all minimal measures in Γ_A as $n \to \infty$ (the strong convergence referring to the metric space $\mathcal{E}/\mathcal{N}$, where $\mathcal{N}$ denotes the class of all measures of zero energy). If k is strictly definite, and hence perfect, the (unique) interior capacitary distribution λ_n associated with A_n converges strongly and vaguely to the interior capacitary distribution associated with A.

Proof of Theorem 4.2. It suffices to consider the case $\lim \text{cap}_* A_n < +\infty$. Like in the proof of Theorem 4.1, one shows that $\{\lambda_n\}$ is a strong Cauchy sequence in $\mathcal{E}^+$, and hence it converges strongly to λ, where λ denotes an arbitrary vague cluster point for the vaguely bounded sequence $\{\lambda_n\}$. Moreover, $\lambda \in \Gamma_{A_p}$ for every p because $\lambda_n \in \Gamma_{A_p}$ for $n > p$. From the corollary to Lemma 2.3.5 follows now easily that $\lambda \in \Gamma_A$. The strong convergence $\lambda_n \to \lambda$ implies further that

$$\|\lambda\|^2 = \lim_n \|\lambda_n\|^2 = \lim \text{cap}_* A_n \leqslant \text{cap}_* A.$$

Hence λ is minimal in Γ_A, and $\text{cap}_* A = \lim \text{cap}_* A_n$. The fact that λ adheres vaguely to $\{\lambda_n\}$ implies in the usual way that $\lambda \geqslant 0$, $V(\lambda) \leqslant 1$, $\lambda(X) = \text{cap}_* A$, and also that $S(\lambda) \subset \bar{A}$ if $S(\lambda_n) \subset \bar{A}_n$. This completes the proof.

LEMMA 4.2.1. *Suppose the kernel* k *is consistent. Consider a decreasing filtering family* $\mathfrak{F}$ *of arbitrary sets* $A \subset X$ *with* $\text{cap}_* A < +\infty$, *and write*

$$A_0 = \bigcap_{A \in \mathfrak{F}} A.$$

If λ_A *denotes an interior capacitary distribution associated with* $A \in \mathfrak{F}$, (2) *then every vague*

(1) This latter part of the theorem may be formulated in a slightly stronger way as follows. The sections

$$S_p = \bigcup_{n > p} \Lambda_{A_n} \qquad (p = 1, 2, \ldots)$$

form the base of a strong Cauchy filter on $\mathcal{E}^+$ whose vague adherence $\bigcap_p \bar{S}_p$ is non-void and contained in Λ_A. A similar result subsists with Λ'_{A_n} and Λ'_A in place of Λ_{A_n} and Λ_A, respectively.

(2) Again one obtains a stronger formulation by introducing the filter Φ on $\mathcal{E}^+$ based on the totality of sections

cluster point λ_0 *along* $\mathfrak{F}$ *of the vaguely bounded family of measures* $\{\lambda_A\}_{A\in\mathfrak{F}}$ *has the properties* (a), (b), *and* (c) *stated in Theorem* 4.1 (but with A_0 in place of A), *and*

$$\|\lambda_0\|^2 = \lambda_0(X) = \inf_{A\in\mathfrak{F}} \operatorname{cap}_* A.$$

Moreover, $\lambda_A \to \lambda_0$ *strongly along* $\mathfrak{F}$. *In the special case where all the sets* $A\in\mathfrak{F}$ *are closed and where* $S(\lambda_A)\subset A$, *we have*

$$\operatorname{cap}_* A_0 = \inf_{A\in\mathfrak{F}} \operatorname{cap}_* A, \tag{1}$$

and λ_0 *is therefore an interior capacitary distribution on* A_0.

Proof. The "sections" $\sum_A = \{\lambda_B : B\in\mathfrak{F},\ B\subset A\}$ constitute the base of a filter Φ for which λ_0 is a vague cluster point. It is easily shown in the usual way that Φ is a strong Cauchy filter. In view of the consistency of the kernel k, we infer that $\Phi\to\lambda_0$ strongly. Hence

$$\|\lambda_0\|^2 = \lim_{A\in\mathfrak{F}} \|\lambda_A\|^2 = \lim_{A\in\mathfrak{F}} \operatorname{cap}_* A = \inf_{A\in\mathfrak{F}} \operatorname{cap}_* A.$$

Since $\lambda_A\in\Gamma_A\subset\Gamma_{A_0}$ for every $A\in\mathfrak{F}$, we obtain $\lambda_0\in\Gamma_{A_0}$. The remaining properties (b), (c), and $\lambda_0(X) = \|\lambda_0\|^2$ are derived in the usual way from the fact that λ_0 adheres vaguely to Φ. Moreover, $S(\lambda_A)\subset\bar{A}$ implies $S(\lambda_0)\subset\bar{A}$, and hence

$$S(\lambda_0)\subset \bigcap_{A\in\mathfrak{F}} \bar{A}.$$

(Note that λ_0 need not be supported by $\bar{A}_0$, and that (1) need not hold for arbitrary sets A, cf. note [1], p. 155.) If the sets $A\in\mathfrak{F}$ are *closed,* λ_0 is actually supported by A_0, and it follows from (1), § 2.5, that

$$\operatorname{cap}_* A_0 \geq 2\lambda_0(X) - \|\lambda_0\|^2 = \inf_{A\in\mathfrak{F}} \operatorname{cap}_* A.$$

This establishes (1), and the proof is complete.

LEMMA 4.2.2. *Suppose k is consistent and X is normal. Every closed set of finite exterior capacity is capacitable, and so is every denumerable union of such sets.*

Note that this lemma is contained in Lemma 4.1.2 if the space X is of class K_σ. In the proof of Lemma 4.2.2 we consider first a single closed set F_0 with $\operatorname{cap}^* F_0 < +\infty$. Denote by G some open set such that $G\supset F_0$, and cap $G < +\infty$. Since X is sup-

---

$$S_A = \bigcup_B \Lambda_B, \qquad \text{resp. } S_A = \bigcup_B \Lambda'_B,$$

where $A, B\in\mathfrak{F}$ and $B\subset A$. Any vague cluster point λ_0 of Φ has then the properties listed above.

posed to be normal, G contains some closed neighbourhood H of F_0. Let $\mathfrak{F}$ denote the decreasing filtering family of *all closed neighbourhoods* $F \subset H$ of F_0. In view of the normalcy of X, F_0 is the intersection of this family $\mathfrak{F}$. According to the preceding lemma,

$$\text{cap}^* F_0 \leqslant \inf_{F \in \mathfrak{F}} \text{cap}_* F = \text{cap}_* F_0,$$

and hence F_0 is capacitable. As to the second part of the lemma we refer to the corresponding part of the proof of Lemma 4.1.2. Like in the analogous case of measures, the limitation to closed sets of *finite* exterior capacity cannot be dispensed with in Lemma 4.2.2. This appears from Ex. 10, § 8.3.

4.3. *The exterior capacitary distributions.* We continue the study of a consistent kernel k on a locally compact space X, but we shall now make the following two additional assumptions concerning X:

(H_1) X is normal.

(H_2) Every open set is of class F_σ. (1)

A topological space X possessing these two properties is called *perfectly normal* (Bourbaki [3], § 4, exerc. 7). Clearly, any metrizable space is perfectly normal, but the converse is false even in the case of a compact space (Bourbaki [3], § 2, exerc. 13). An equivalent form of (H_2) is that every closed set should be of class G_δ. Hence any closed subset of a perfectly normal space is representable as the intersection of a sequence of closed neighbourhoods of F; or, in other words, $F = \bigcap \bar{G}_n$, where each G_n is open and contains F. In the case of a metric space with the distance d we may, for instance, use the following open sets:

$$G_n = \{x \in X: \ d(x, F) < n^{-1}\}.$$

The following consequence of (H_2) is of importance in the sequel: Any set of class $(FG)_\sigma$, that is of the form

$$H = \bigcup_{n=1}^{\infty} (F_n \cap G_n) \quad (F_n \text{ closed, } G_n \text{ open}),$$

is actually of class F_σ.

We consider now a consistent kernel k on a locally compact space X satisfying the hypotheses (H_1), (H_2), and we assume as usual that $k \geqslant 0$ unless X is compact.

(1) A set is said to be of class F_σ if it is representable as the union of a sequence of closed sets. A set is of class G_δ if it is representable as the intersection of a sequence of open sets. The classes $F_{\sigma\delta}$ and $G_{\delta\sigma}$, etc., are defined correspondingly.

For brevity, we shall call a set A *σ-compact*, resp. *σ-finite*, if A can be covered by a sequence of sets each of which is compact, resp. of finite exterior capacity. (1) In the latter case, the covering sets may clearly be chosen as open sets, or, in view of (H_2), as *closed* sets. In other words, a set A is σ-finite if and only if A is contained in some set of the type considered in the second part of Lemma 4.2.2. In particular, *every σ-finite* (or σ-compact) *set of class* $(FG)_\sigma = F_\sigma$ *is capacitable.* In view of this observation, the relations between potentials and interior capacity, described in § 3.2 in the case of an arbitrary definite kernel, have exact analogues for the *exterior* capacity under the present assumptions of a consistent kernel on a locally compact space X satisfying the hypotheses (H_1), (H_2).

LEMMA 4.3.1. *Two measures $\lambda, \mu \in \mathcal{E}$ are equivalent if and only if their potentials coincide quasi-everywhere.*

Proof. Suppose $\|\lambda - \mu\| = 0$. According to the corollary to Lemma 2.3.5 it suffices to prove that any set A such that $k(x, \lambda - \mu) > 1/n$ quasi-everywhere in A is of zero exterior capacity, and this is implied by the following lemma.

LEMMA 4.3.2. *Let $0 < t \leqslant +\infty$, $\mu \in \mathcal{E}$, and let A denote a set such that $k(x, \mu) \geqslant t$ quasi-everywhere in A. Then*

$$\text{cap}^* A \leqslant t^{-2} \|\mu\|^2.$$

Proof. In view of the corollary to Lemma 2.3.5 and the fact that $k(x, \mu)$ is defined quasi-everywhere (Lemma 3.2.3), we may assume that $k(x, \mu)$ is defined and $\geqslant t$ *everywhere* in A. For any number s, $0 < s < t$, A is then contained in the set

$$H = \{x \in X : k(x, \mu^+) > k(x, \mu^-) + s\}.$$

Introducing the open sets G_r and the closed sets F_r (r rational) as follows:

$$F_r = \{x \in X : k(x, \mu^-) \leqslant r\},$$
$$G_r = \{x \in X : k(x, \mu^+) > r + s\},$$

we find that H is of class $(FG)_\sigma$, being the union of the sets $F_r \cap G_r$. Hence H is capacitable if σ-finite or σ-compact. It suffices to consider case I ($k \geqslant 0$). Here H is contained in the open set

$$\{x \in X : k(x, \mu^+) > s\},$$

(1) Recall that any σ-compact set is σ-finite provided the kernel k is strictly pseudo-positive (thus in particular if k is strictly positive or strictly definite).

whose capacity is finite according to Lemma 3.2.2. Consequently, H is capacitable, and we infer from Lemma 3.2.2 that

$$\text{cap}^* H = \text{cap}_* H \leqslant s^{-2} \cdot \|\mu\|^2,$$

from which the assertion of the lemma follows for $s \to t$.[1]

LEMMA 4.3.3. *If* $\mu_n \to \mu$ *strongly in* $\mathcal{E}$, *then*

$$k(x, \mu) \geqslant \liminf_n k(x, \mu_n) \qquad \text{quasi-everywhere in } X.$$

This follows from the preceding lemma in the same way as Lemma 3.2.4 followed from Lemma 3.2.3. Again, one may easily prove that there is a subsequence $\{\mu_{n_i}\}$ whose potentials converge quasi-everywhere to the potential $k(x, \mu)$ of μ. Cf. J. Deny [15], Théorème 1, c).

With any given set $A \subset X$ we associate the convex class

$$\Gamma_A^* = \{\mu \in \mathcal{E} : k(x, \mu) \geqslant 1 \text{ quasi-everywhere in } A\}.$$

It follows immediately from the last lemma above that Γ_A^* is sequentially strongly closed in $\mathcal{E}$. Since the strong topology on $\mathcal{E}$ is defined by means of a semi-norm, Γ_A^* is actually closed in the strong topology. According to Lemma 4.3.2 with $t = 1$, $\|\mu\|^2 \geqslant \text{cap}^* A$ for every $\mu \in \Gamma_A^*$; that is, $\|\Gamma_A^*\|^2 \geqslant \text{cap}^* A$ (cf. Lemma 4.1.1). In particular, Γ_A^* is void unless $\text{cap}^* A < +\infty$. We proceed to prove that, actually,

$$\|\Gamma_A^*\|^2 = \text{cap}^* A.$$

It suffices to establish the existence of a single measure $\lambda \in \Gamma_A^*$ of energy $\text{cap}^* A$ under the assumption $\text{cap}^* A < +\infty$; and this will be done in the following theorem.

THEOREM 4.3. *Suppose that the kernel k is consistent and that the locally compact space X is perfectly normal. For any set* $A \subset X$ *with* $\text{cap}^* A < +\infty$, *the class* Λ_A^* *of all positive measures* λ *for which*

$$\|\lambda\|^2 = \lambda(X) = \text{cap}^* A$$

[1] This lemma is due to H. Cartan [10], Lemme 5, p. 98, in the Newtonian case; but the manner in which Cartan reduced it to the elementary Lemma 3.2.2 was different from the procedure employed above. Cartan's proof depends on the fact that the Newtonian kernel has the property described in the second part of Theorem 3.4.1. His method is, therefore, applicable to any *regular* kernel (even without restrictions on the locally compact space X).

and (a) $k(x, \lambda) \geqslant 1$ *quasi-everywhere in* A,

(b) $k(x, \lambda) \leqslant 1$ *everywhere in the support of* λ,

and hence (c) $k(x, \lambda) = 1$ λ*-almost everywhere in* X,

is non-void and vaguely compact.

The same is true of the class $\Lambda_A^{*\prime}$ of all measures $\lambda \in \Lambda_A^*$ supported by $\bar{A}$. The measures of class Λ_A^* are called the *exterior capacitary distributions* associated with the set A. The remarks to Theorem 4.1 have obvious analogues in the present case.

Proof of Theorem 4.3. In the special case of an *open* set $A = G$, any interior capacitary distribution λ associated with G has the stated properties because the set

$$H = G \cap \{x \in X : \; k(x, \lambda) < 1\}$$

is of class $GF_\sigma = F_\sigma$ and hence capacitable (since $\text{cap}^* H \leqslant \text{cap } G < +\infty$). To an *arbitrary* set A with $\text{cap}^* A < +\infty$ corresponds a sequence $\{G_n\}$ of open sets containing A such that $\text{cap } G_n < +\infty$ and

$$\lim_n \text{cap } G_n = \text{cap}^* A.$$

Since X is perfectly normal, we may assume that $\bigcap_n \bar{G}_n = \bar{A}$. Replacing, if necessary, G_n by $G_1 \cap G_2 \cap \ldots \cap G_n$, we may suppose, further, that $G_n \subset G_p$ for $n > p$. Denoting by λ_n an interior capacitary distribution associated with G_n and supported by $\bar{G}_n$, we conclude in the usual way that $\{\lambda_n\}$ is a strong Cauchy sequence in $\mathcal{E}^+$. Since $\lambda_n(X) = \text{cap } G_n$ remains bounded, there exists a vague cluster point λ for $\{\lambda_n\}$, and hence $\lambda_n \to \lambda$ strongly in view of the consistency of k. Clearly $\lambda_n \in \Gamma_{G_n}^* \subset \Gamma_A^*$, and hence $\lambda \in \Gamma_A^*$. Moreover, λ_n is supported by $\bar{G}_p$ for $n > p$, and hence λ is supported by each $\bar{G}_p$, i.e.,

$$S(\lambda) \subset \bigcap_p \bar{G}_p = \bar{A}.$$

Proceeding as in the proof of Theorem 4.1, we conclude that $\lambda \in \Lambda_A^{*\prime}$. Finally, it is shown in the previous way that Λ_A^* and $\Lambda_A^{*\prime}$, which are obviously vaguely bounded, are vaguely closed and hence compact.

LEMMA 4.3.4 *If* A *is capacitable and of finite capacity,* $\Lambda_A = \Lambda_A^*$ *and* $\Lambda_A' = \Lambda_A^{*\prime}$.

In fact, the minimal measures in Γ_A and in Γ_A^* constitute two equivalence classes in $\mathcal{E}$ of which the former contains the latter because $\Gamma_A \supset \Gamma_A^*$ and $\|\Gamma_A\|^2 = \|\Gamma_A^*\|^2$ ($= \text{cap } A$). Consequently, these two equivalence classes are identical, and the lemma follows easily.

4.4. *Monotone families of sets, and the associated exterior capacitary distributions.* The results in § 4.2 concerning monotone families of sets can be extended with unchanged proofs to the case of exterior capacity and exterior capacitary distributions. The limitation to sets of class $\mathfrak{A}$ in Theorem 4.2 is unnecessary here on account of the corollary to Lemma 2.3.5. Particularly interesting is the following theorem, which corresponds to the first part of Theorem 4.2, and which, in the terminology of Choquet [14], § 15.3, asserts that the capacity associated with the kernel k is *alternating* of order 1, a.

THEOREM 4.4. *Suppose that the kernel k is consistent and that the locally compact space X is perfectly normal. If A denotes the union of an increasing sequence of arbitrary sets $A_n \subset X$, then*

$$\text{cap}^* A = \lim_n \text{cap}^* A_n.$$

COROLLARY. Any denumerable union of capacitable sets of class $\mathfrak{A}$ is capacitable.

4.5. *Application of Choquet's theory of capacitability.*

THEOREM 4.5. *Suppose that the kernel k is consistent and that the locally compact space X is perfectly normal. Every K-analytic set is capacitable, and so is every σ-compact or σ-finite Borel set.*

The hypothesis (H_2) implies that the different classical definitions of a Borel set are, in fact, equivalent (but the concept of a Borel set is more general than that of a K-borelian set in the sense of Choquet [14] unless X is of class K_σ). However, any *σ-compact* Borel set is a K-borelian set, and the capacitability of such a set follows, therefore (in view of Theorem 4.4 above), from part (i) of a fundamental theorem in Choquet [14] (Théorème 30.1). Part (ii) of the same theorem shows that any *K-analytic* subset of X is capacitable (likewise on account of our Theorem 4.4). We shall omit the proof of the capacitability of all *σ-finite* Borel sets because this proof depends on various adaptations of Choquet's theory to the present circumstances. Of course, this last case of σ-finite Borel sets is of interest only if X is not of class K_σ. The result should be compared with the well-known fact from measure theory that $\mu^*(A) = \mu_*(A)$ if A is μ-measurable and contained in the union of a sequence of μ-integrable sets. In both cases, the σ-finiteness condition on A cannot be dispensed with (cf. Example 10, § 8.3).

5. Extensions of the theory

5.1. *The case of a continuous weight function.* The preceding theory can be extended when we replace the function 1 in the properties (a), (b), and (c) of the capacitary potentials (cf. Theorems 2.5, 4.1, and 4.3) by a more general function $f \geqslant 0$, and hence $\mu(X) = \int 1\, d\mu$ by $\int f\, d\mu$, etc. The simplest case is that of a continuous function with values $0 < f(x) < +\infty$. The preceding theory can be carried over to this new case by obvious modifications in the proofs. Alternatively, one may reduce the new case to the old case by introducing the kernel

$$k_q(x, y) = k(x, y)\, q(x)\, q(y), \qquad q = 1/f.$$

The kernels k and k_q are always of the same kind (say, positive, pseudo-positive, symmetric, definite, strictly definite, consistent, or perfect). In fact, the mapping

$$\mu \to q \cdot \mu$$

is a linear homeomorphism of $\mathfrak{M}(X)$ (with the vague topology) onto itself, leaving $\mathfrak{M}^+(X)$ invariant. The identity

$$k_q(\mu, \nu) = k(q\mu, q\nu)$$

shows that, in case of a definite kernel k, the above mapping carries the space $\mathcal{E}_q$ of all measures of finite energy with respect to k_q isometrically onto the space $\mathcal{E}$ formed with respect to k. Note also that the potentials are related as follows:

$$k_q(x, \mu) = q(x) \cdot k(x, q\mu).$$

In particular, the conditions stated in Lemmas 3.4.2 and 3.4.3 are satisfied simultaneously by k and k_q. Likewise, k and k_q are simultaneously regular.

5.2. *Balayage.* Suppose now the kernel k is *definite*, and take for f the potential $f(x) = k(x, \omega)$ of a given measure ω of finite energy with respect to k. (The assumption $f \geqslant 0$ is fulfilled at least if $\omega \geqslant 0$ and $k \geqslant 0$, but f is now in general discontinuous.) Replacing $\mu(X)$ by $\int f\, d\mu = k(\mu, \omega)$ in the definition (1), § 2.5, of the interior Wiener capacity, we define for any set $A \subset X$

$$\operatorname{cap}_{\omega *} A = \sup_{\lambda} \{2k(\lambda, \omega) - k(\lambda, \lambda)\} = \|\omega\|^2 - \inf_{\lambda} \|\lambda - \omega\|^2,$$

where λ ranges over the class of all positive measures of finite energy and concentrated on A (or supported by some compact subset of A). Clearly,

$$0 \leqslant \operatorname{cap}_{\omega *} A \leqslant \operatorname{cap}_{\omega *} X < +\infty.$$

We shall now make the assumption that the space $\mathcal{E}_F^+$ of all positive measures of finite energy supported by F is strongly *complete* for every closed set F. (Any perfect kernel fulfills this assumption, and so does any consistent and strictly pseudo-positive kernel, cf. § 3.3.) In view of this assumption the above extremal problems have solutions when $A = F$ is closed. In fact the problem is to minimize the distance between ω and the points λ of a convex, complete subset $\mathcal{E}_F^+$ of a pre-Hilbert space $\mathcal{E}$. The process of solution of this problem (the "projection" of ω onto $\mathcal{E}_F^+$) is called the "sweeping-out" process (French: "balayage"). The resulting measures λ are characterized within $\mathcal{E}_F^+$ by the relations

$$\|\lambda\|^2 = k(\lambda, \omega) = \operatorname{cap}_{\omega *} F,$$

and their potentials have the following properties:

(a) $k(x, \lambda) \geqslant k(x, \omega)$ in F except in some subset N with $\operatorname{cap}_{\omega *} N = 0$,

(c) $k(x, \lambda) = k(x, \omega)$ λ-almost everywhere in X.

There seems to be no obvious analogue of property (b) in Theorem 2.5 (unless $f = k(x, \omega)$ is continuous and > 0, in which case we are back in the case considered above in § 5.1). This is because all reference to the vague topology has disappeared.

Next one may define an exterior capacity $\operatorname{cap}_\omega^* A = \inf \operatorname{cap}_{\omega *} G$ as G ranges over the class of all open sets containing A. Open sets and closed sets are capacitable. If $k \geqslant 0$, this exterior capacity is countably subadditive. A theory of *interior balayage* can be developed in analogy with §§ 4.1, 4.2 (again except for property (b)). There is a similar theory of *exterior balayage* under the additional hypotheses (H_1), (H_2), § 4.3. Finally, it can be shown as before that all K-analytic sets and all Borel sets are capacitable.

5.3. *The case of a kernel of variable sign on a non-compact space.* So far, we have studied the two cases I: $k \geqslant 0$, and II: X compact. In the general case of a kernel k of variable sign on a locally compact, not necessarily compact space X, we shall limit the attention to measures of (uniformly) compact supports. As mentioned in § 2.1, this case can be reduced to case II simply by replacing the kernel k by its restriction to $K \times K$, where K denotes an arbitrary compact subset of X. It is, however, sometimes preferable to remain in the original locally compact space X; and hence the following definitions are convenient (and sometimes of interest even in Case I of a positive kernel, cf. Theorem 3.4.2):

A kernel k on a locally compact space X is called *K-definite, strictly K-definite, K-consistent,* or *K-perfect* if, for every compact set $K \subset X$, the restriction of k to

$K \times K$ is definite, strictly definite, consistent, or perfect, respectively. In Case I (or Case II), *a kernel k is definite if and only if it is K-definite.* This follows from Lemma 2.2.2 and the remark thereto, applied to the three terms in the decomposition

$$k(\mu, \mu) = k(\mu^+, \mu^+) + k(\mu^-, \mu^-) - 2k(\mu^+, \mu^-).$$

A kernel which is strictly definite, consistent, or perfect (in Case I or Case II) is likewise strictly K-definite, K-consistent, or K-perfect, respectively. The converse is trivial in Case II and false in Case I.

Consider now a *K-definite* kernel k on a locally compact space X. (According to Theorem 3.3, such a kernel is K-perfect if and only if it is K-consistent and strictly K-definite.) We denote by $\mathcal{E}_K$, resp. $\mathcal{E}_K^+$, the space of all measures, resp. positive measures, of finite energy and supported by the compact set $K \subset X$. It follows then

(a) from Lemma 3.3.1 that $\mathcal{E}_K^+$ is strongly complete for every compact set $K \subset X$, provided k is K-consistent and strictly pseudo-positive.

(b) from the proof of Lemma 3.4.2 that k is K-consistent if there corresponds to every compact set $K \subset X$ a compact set $K_1 \supset K$ with the property that every measure $\mu \in \mathcal{E}_K$ can be approximated in the strong sense by measures $\lambda \in \mathcal{E}_{K_1}$ for which the potential $k(x, \lambda)$ is continuous (or at least: the restriction of $k(x, \lambda)$ to K is continuous).[1]

(c) from Theorem 3.4.1 that k is K-consistent if it is regular.

Let us now suppose that k is *K-consistent* (cf. Theorem 3.4.2). From the results of § 4 follows that there corresponds to every *relatively compact* set $A \subset X$ of finite interior capacity a non-void, vaguely compact class of interior capacitary distributions (some of which are supported by $\bar{A}$). A similar result holds for the exterior capa-

---

[1] In order to verify that the restriction of k to $K \times K$ has the property (CW) formulated in Lemma 3.4.1, we consider a filter Φ on the part of $\mathcal{E}_K^+$ determined by $\|\mu\| \leq M$, and suppose that Φ converges vaguely to some measure μ_0. It follows that $\mu_0 \in \mathcal{E}_K^+$ and $\|\mu_0\| \leq M$. To any number $\eta > 0$ and any measure $\lambda \in \mathcal{E}_K^+$ corresponds, by assumption, a measure $\lambda' \in \mathcal{E}_{K_1}^+$ such that $\|\lambda - \lambda'\| < \eta$ and the restriction of $k(x, \lambda')$ to K is continuous. The vague convergence $\Phi \to \mu_0$ implies

$$k(\mu, \lambda') = \int k(x, \lambda')\, d\mu \to \int k(x, \lambda')\, d\mu_0 = k(\mu_0, \lambda')$$

along Φ. Using the Cauchy-Schwarz inequality, we obtain

$$|k(\mu - \mu_0, \lambda)| \leq |k(\mu - \mu_0, \lambda')| + \|\mu - \mu_0\| \; \|\lambda - \lambda'\|,$$

and hence, along Φ, $\limsup_{\mu} |k(\mu, \lambda) - k(\mu_0, \lambda)| \leq 2M\eta$ because $\|\mu - \mu_0\| \leq \|\mu\| + \|\mu_0\| \leq 2M$. Letting $\eta \to 0$, we conclude that Φ converges weakly to μ_0.

citary distributions provided the space X has the property that every relatively compact, open set is of class K_σ (thus in particular if every compact subset of X is metrizable). Finally, every relatively compact Borel set is capacitable, and so is every relatively compact K-analytic set.

III. Convolution Kernels

We shall now study the important case where X is a locally compact topological *group* and where the kernel $k(x, y)$ is *invariant* under, say, right multiplication:

$$k(xz, yz) = k(x, y) \qquad (x, y, z \in X).$$

Writing briefly $k(x)$ in place of $k(x, e)$, where e denotes the unit element in the group X, we find the expression

$$k(x, y) = k(xy^{-1})$$

for the kernel $k(x, y)$ in terms of the *kernel function* $k = k(x)$. The potential of a measure μ with respect to a (right) invariant kernel is, therefore, the convolution $k \ast \mu$ of the kernel function k and the measure μ (cf. § 6.2). For this reason we shall use the term *convolution kernel* synonymously with right invariant kernel on a locally compact group.

6. Preliminaries concerning locally compact topological groups

The following notations are convenient. The *reflex* of a (real-valued) function f on the group X is the function $\check{f}$ defined by

$$\check{f}(x) = f(x^{-1}).$$

The reflex of a measure μ is the measure $\check{\mu}$ defined by $d\check{\mu}(x) = d\mu(x^{-1})$, i.e.

$$\int f\, d\check{\mu} = \int \check{f}\, d\mu$$

for every $f \in C_0(X)$, and hence also for any function f such that one of the two integrals (and hence also the other) is defined. A function, or a measure, is called *symmetric* if it coincides with its reflex. The right translates of a function f are the functions $x \to f(xa)$, $a \in X$. The right translates of a measure μ are the measures $f \to \int f(xa)\, d\mu(x)$, $a \in X$. Left translates of functions and measures are defined correspondingly.

6.1. *Convolution of two measures* (Cf. H. Cartan [9], § II). Consider first two positive measures μ and ν on the locally compact group X. The mapping

$$f \to \iint f(xy)\, d\mu(x)\, d\nu(y) \qquad (f \in C_0^+(X))$$

is an additive and positive-homogeneous functional on C_0^+ with values $\geqslant 0$ and $\leqslant +\infty$. *If the double integral is finite for every* $f \in C_0^+$, it can be extended in a unique and obvious way to a positive, linear functional on $C_0(X)$, and hence there is a uniquely determined measure $\mu * \nu$, called the convolution of μ and ν, such that the equation

$$\int f \cdot d(\mu * \nu) = \iint f(xy)\, d\mu(x)\, d\nu(y) \tag{1}$$

holds for every function $f \in C_0(X)$. By application of Théorème 1 in Bourbaki [4], Chap. IV, § 1, we infer that (1) holds for a lower semi-continuous function f on X provided either $f \geqslant 0$ or μ and ν have compact supports. (The latter case is reduced to the former in the usual way by adding to f some constant $c \geqslant 0$ such that $f(x)+c \geqslant 0$ in the set $S(\mu)\cdot S(\nu)$.)

For two measures of variable sign one defines the convolution by

$$\mu * \nu = \mu^+ * \nu^+ + \mu^- * \nu^- - \mu^+ * \nu^- - \mu^- * \nu^+,$$

provided the four convolutions on the right are well defined according to the above definition. It is well known that this is the case, in particular, if at least one of the measures μ and ν is of compact support. (1)

The convolution product is commutative if and only if the group is Abelian. In any case the identity

$$(\mu * \nu)^{\vee} = \check{\nu} * \check{\mu} \tag{2}$$

holds in the sense that both convolutions are simultaneously defined or undefined. Each of the distributive laws

$$\lambda * (\mu \pm \nu) = \lambda * \mu \pm \lambda * \nu; \qquad (\mu \pm \nu) * \lambda = \mu * \lambda \pm \nu * \lambda,$$

subsists provided the two convolutions on the right are defined.

As usual, we denote by ε_x the mass $+1$ placed at the point $x \in X$. For ε_e we write simply ε. Clearly, ε is the identity for the convolution product of measures:

$$\mu * \varepsilon = \varepsilon * \mu = \mu.$$

More generally, $\mu * \varepsilon_a$ and $\varepsilon_a * \mu$ are the right and left translates of μ.

(1) In fact, if $\mu \geqslant 0$, $\nu \geqslant 0$, and if μ has compact support, then the supports of $f(xy)$ and of $\mu \otimes \nu$ have a compact intersection in $X \times X$ for any given $f \in C_0(X)$.

6.2. *Convolution of functions and measures.* Consider a lower semi-continuous function f and a measure μ. Under the assumption that either $f \geqslant 0$ or $S(\mu)$ is compact, we define the convolutions $f * \mu$ and $\mu * f$ as functions of x by

$$(f * \mu)(x) = \int f(x y^{-1})\, d\mu(y),$$

$$(\mu * f)(x) = \int f(y^{-1} x)\, d\mu(y),$$

at any point $x \in X$ for which the integral in question is defined (cf. § 1.1). If $\mu \geqslant 0$, then $f * \mu$ and $\mu * f$ are defined and lower semi-continuous everywhere in X. It is easily verified that

$$(f * \mu)\check{} = \check{\mu} * \check{f} \tag{1}$$

in the sense that both expressions are defined at the same points of X. If ν denotes another measure, likewise of compact support unless $f \geqslant 0$, it follows from Fubini's theorem (§ 1.1) that

$$\mu * (f * \nu) = (\mu * f) * \nu = \mu * f * \nu \tag{2}$$

at any point $x \in X$ for which the last expression exists according to the following *definition* as a double integral:

$$(\mu * f * \nu)(x) = \iint f(y^{-1} x z^{-1})\, d\mu(y)\, d\nu(z). \tag{3}$$

This is the case, in particular, if $\mu \geqslant 0$ and $\nu \geqslant 0$. Finally, it follows from Fubini's theorem that the relations

$$\begin{aligned} f * (\mu * \nu) &= (f * \mu) * \nu, \\ (\mu * \nu) * f &= \mu * (\nu * f), \end{aligned} \tag{4}$$

hold, say by positive measures μ and ν, in each of the following two cases: a) $f \geqslant 0$ and $\mu * \nu$ exists; b) μ and ν have compact supports.

6.3. *The Haar measures.* It is well known that there exist on any locally compact topological group X positive measures $(\neq 0)$ which are invariant under right or left translations. These invariant measures are called the *Haar measures* on X. It is also known that, if m denotes some right invariant Haar measure, then any other right invariant measure on X is a constant multiple of m. Similarly in case of left invariant measures. If m is right invariant, $\check{m}$ is left invariant. If X is Abelian or compact, $\check{m} = m$, so that any right invariant measure is left invariant, and conversely. In the general case, there is a certain continuous function $\varrho = \varrho(x) > 0$ on X, the *modular function*, with the property that

$$\check{m} = \varrho \cdot m.$$

Moreover, $\varrho(xy) = \varrho(x)\varrho(y)$; $\varrho(x^{-1}) = 1/\varrho(x)$; $\varrho(e) = 1$.

In the sequel we shall let m stand for a fixed *right* invariant Haar measure on the group X.

To any locally m-integrable function f on X corresponds a measure fm with the density f. For such a function f the convolution concept defined in § 6.2 may be reduced to that of § 6.1. In fact, let f denote a lower semi-continuous, locally m-integrable function, and let μ be a positive measure (of compact support unless $f \geqslant 0$). Then the convolution $(fm) * \mu$ of the measures fm and μ is well-defined if and only if $f * \mu$ is locally m-integrable. In the affirmative case (e.g., if $S(\mu)$ is compact), the measure $(fm) * \mu$ has the density $f * \mu$:

$$(fm) * \mu = (f * \mu) \cdot m. \tag{1}$$

In fact, for any function $\varphi \in C_0^+(X)$,

$$\int \varphi\, d((f * \mu) m) = \int \varphi(x) \left\{ \int f(xy^{-1})\, d\mu(y) \right\} dm(x) = \int \left\{ \int \varphi(x) f(xy^{-1})\, dm(x) \right\} d\mu(y)$$
$$= \int \left\{ \int \varphi(zy) f(z)\, dm(z) \right\} d\mu(y),$$

which equals the integral of φ with respect to $(fm) * \mu$. Similarly,

$$\mu * (f\check{m}) = (\mu * f) \cdot \check{m}. \tag{2}$$

The convolution of *two functions* f and g will be defined by

$$(f * g)(x) = \int f(xy^{-1})\, g(y)\, dm(y) = \int f(y)\, g(y^{-1}x)\, d\check{m}(y). \tag{3}$$

We shall only need the case where f and g are lower semi-continuous and $\geqslant 0$. Then $f * g$ is likewise lower semi-continuous (since $f(xy^{-1})\, g(y)$ is lower semi-continuous on $X \times X$, and $m \in \mathfrak{M}^+(X)$). If g is locally Haar integrable, $f * g = f * (gm)$; if f is locally Haar integrable, $f * g = (f\check{m}) * g$. The following identities are easily verified:

$$(f * g)^{\vee} = \check{g} * \check{f}, \tag{4}$$

$$\int fg\, dm = (\check{f} * g)(e) = (\check{g} * f)(e), \tag{5}$$

$$\int f\, d\mu = (\check{f} * \mu)(e) = (\check{\mu} * f)(e). \tag{6}$$

6.4. *Kernel function. Energy function.* A lower semi-continuous function $k = k(x)$ on the locally compact group X will be called a *kernel function* on X. The corresponding convolution kernel is defined by $k(x, y) = k(xy^{-1})$. This kernel is clearly positive or symmetric if and only if the kernel function k is positive ($k \geq 0$) or symmetric ($\check{k} = k$). In a similar way we shall call the kernel function definite, consistent, etc., if the kernel $k(xy^{-1})$ is definite, consistent, etc.

The potential of a measure μ (of compact support unless $k \geq 0$) is simply the convolution $k * \mu$. The mutual energy $k(\mu, \nu)$ of two measures μ and ν (of compact support unless $k \geq 0$) is the value of $\check{\mu} * k * \nu$ (cf. (3), § 6.2) at the group identity e:

$$k(\mu, \nu) = (\check{\mu} * k * \nu)(e). \tag{1}$$

In particular, the energy $k(\lambda, \lambda)$ of a measure λ (of compact support unless $k \geq 0$) is the value $k_\lambda(e)$ at e of the so-called *energy function*

$$\begin{aligned} k_\lambda &= \check{\lambda} * k * \lambda, \\ k_\lambda(x) &= \iint k(s x t^{-1})\, d\lambda(s)\, d\lambda(t), \end{aligned} \tag{2}$$

associated with the kernel function k (cf. J. Deny [16]).

Lemma 6.4.1. *For any kernel function k and any measure $\lambda \geq 0$* (of compact support unless $k \geq 0$), *the energy function $k_\lambda = \check{\lambda} * k * \lambda$ is itself a kernel function on the group X. Whenever defined, the corresponding mutual energy of two measures μ and ν of compact support is determined by*

$$k_\lambda(\mu, \nu) = k(\lambda * \mu, \lambda * \nu). \tag{3}$$

If k is symmetric or K-definite, k_λ has the same property.

Proof. The only point requiring a comment is the validity of (3). The case of measures μ and ν of variable sign is easily reduced to that of positive measures, and here (3) follows from (1) by repeated applications of the associative laws (2), (4), § 6.2, and the identity (2), § 6.1:

$$\check{\mu} * k_\lambda * \nu = \check{\mu} * (\check{\lambda} * k * \lambda) * \nu = (\lambda * \mu)^\vee * k * (\lambda * \nu). \tag{4}$$

7. Definite convolution kernels

In the present section we assume that the lower semi-continuous function $k = k(x)$ (the kernel function) on the locally compact group X is symmetric ($\check{k} = k$) and *K-definite*. Thus

$$\|\mu\|^2 = k(\mu, \mu) = (\check{\mu} * k * \mu)(e) \geq 0$$

for every measure μ of compact support for which the energy $k(\mu, \mu)$ is defined (i.e., $k(\mu^+, \mu^+) + k(\mu^-, \mu^-)$ and $k(\mu^+, \mu^-)$ are not both infinite, cf. § 2.1). As pointed out in § 5.3, this implies that k is *definite* provided either $k \geqslant 0$ or X is compact.

7.1. *Continuity of the energy function.* The key to the discussion of K-definite convolution kernels is the following result concerning real-valued positive definite functions (in the usual sense) on a quite arbitrary topological group: (1)

LEMMA 7.1.1. *Let f denote a lower semi-continuous, symmetric function on a topological group X, with values $-\infty < f(x) \leqslant +\infty$. Suppose f is positive definite in the sense that*

$$\sum_{i,j} f(x_i x_j^{-1})\, q_i q_j \geqslant 0 \tag{1}$$

for any finite set of points $x_j \in X$ and finite real numbers q_i such that the sum on the left is meaningful. If, in addition, the value $f(e)$ of f at the unit element of X is finite, then f is bounded and uniformly continuous.

Proof. Applying (1) to the single point e with the weight $q = 1$, we get $f(e) \geqslant 0$. Using two points e and x with weights 1 and q, we obtain

$$f(e) + 2f(x)\,q + f(e)\,q^2 \geqslant 0.$$

The sum on the left is always meaningful because $f(e)$ is finite. It follows that $f(x)$ is finite. Hence the determinant $f(e)^2 - f(x)^2$ is $\geqslant 0$, i.e.,

$$|f(x)| \leqslant f(e). \tag{2}$$

Since $f(e) = 0$ implies $f = 0$, we assume in the sequel that $f(e) \neq 0$. Another consequence of (2) is that the lower semi-continuous function f is continuous at the point e:

$$f(e) \leqslant \liminf_{x \to e} f(x) \leqslant \limsup_{x \to e} f(x) \leqslant f(e).$$

Applying (1) with three points, e, x, y, we obtain after evaluating the determinant D of the quadratic form in q_1, q_2, q_3 on the left:

$$0 \leqslant D \cdot f(e) = (f(e)^2 - f(x)^2)(f(e)^2 - f(y)^2) - (f(e) f(xy^{-1}) - f(x) f(y))^2.$$

Hence, $$|f(e) f(xy^{-1}) - f(x) f(y)| \leqslant f(e) \cdot (f(e)^2 - f(y)^2)^{\frac{1}{2}}.$$

(1) This lemma shows that Bochner's theorem on spectral representation of continuous, positive definite functions, say on the real line, remains valid if the assumption of continuity is replaced by that of lower semi-continuity.

Now, $$f(e)\,|f(xy^{-1})-f(x)|\leqslant|f(e)f(xy^{-1})-f(x)f(y)|+|f(x)|\cdot|f(y)-f(e)|.$$

Inserting $|f(x)|\leqslant f(e)$, and combining with the preceding inequality, we obtain

$$|f(xy^{-1})-f(x)|\leqslant(f(e)^2-f(y)^2)^{\frac{1}{2}}+|f(e)-f(y)|. \tag{3}$$

Replacing x and y by x^{-1} and y^{-1}, respectively, and using the symmetry of f, we get the same estimate for $|f(y^{-1}x)-f(x)|$. The uniform continuity follows now from the fact that the right hand side of (3) is independent of x and approaches 0 as $y\to e$.

Returning to the study of convolution kernels on a locally compact topological group X, we infer immediately from this lemma that any K-definite kernel function k on X with $k(e)\neq+\infty$ is bounded and uniformly continuous. In fact, such a kernel function is, by definition, symmetric, and it is positive definite in the sense (1) because the left hand side of (1) is the energy of the measure $\mu=\sum_i q_i\,\varepsilon_{x_i}$, the support of which is finite and hence compact. Actually, most kernel functions of interest are unbounded: $k(e)=+\infty$. A more interesting application of Lemma 7.1.1 is described in the following theorem.

THEOREM 7.1. *Let k denote a K-definite kernel function on a locally compact group X. For any measure λ of finite energy* (and of compact support unless $k\geqslant 0$), *the energy function*

$$k_\lambda=\check{\lambda}*k*\lambda$$

is everywhere defined, bounded, uniformly continuous, and K-definite.

Proof. In the case $\lambda\geqslant 0$ it follows from Lemma 6.4.1 that k_λ is itself a K-definite kernel function on X. Since $k_\lambda(e)=\|\lambda\|^2<+\infty$, we conclude from Lemma 7.1.1, in view of the above observation, that k_λ is bounded und uniformly continuous. In the general case we apply this result to the positive measures λ^+, λ^-, and $|\lambda|=\lambda^++\lambda^-$, each of which is of finite energy (and of compact support unless $k\geqslant 0$). The consideration of $|\lambda|$ shows that k_λ is everywhere defined and finite. In view of the "parallellogram law" for symmetric bilinear forms,

$$k_\lambda=2\,k_{\lambda^+}+2\,k_{\lambda^-}-k_{|\lambda|}. \tag{4}$$

This formula shows that k_λ is bounded and uniformly continuous. Applying (3), § 6.4, with λ replaced by λ^+, λ^-, or $|\lambda|$, and with $\nu=\mu$, we obtain in view of (4)

$$k_\lambda(\mu,\mu)=\|\lambda*\mu\|^2$$

for any measure μ of compact support. This shows that k_λ is K-definite.

The following remark connected with Theorem 7.1 will not be used in the sequel:

Remark. Suppose, in addition, that the group X is *unimodular* ($\varrho = 1$, $\check{m} = m$), and that the kernel function $k = k(x)$ is positive ($k \geqslant 0$) and *locally Haar integrable* (cf. Lemma 7.2.2 below). The measure $\varkappa = k \cdot m$ of density k with respect to Haar measure is then symmetric. We call it the "kernel measure" associated with the kernel function k. For any measure $\lambda \geqslant 0$ the identity

$$\check{\lambda} * \varkappa * \lambda = k_\lambda \cdot m \tag{5}$$

holds in the sense that the convolution on the left is defined (1) if and only if k_λ is locally Haar integrable (cf. the proof of the analogous identity (1), § 6.3). We shall now establish the following result: (2)

A necessary and sufficient condition that $\lambda \geqslant 0$ be of finite energy with respect to $k(xy^{-1})$ is that the measure $\check{\lambda} * \varkappa * \lambda$ exist and have a continuous density with respect to Haar measure. In the affirmative case, the energy of λ equals the value of this density at the group identity e.

The necessity follows immediately from Theorem 7.1 and the above formula (5). As to the sufficiency, suppose $\check{\lambda} * \varkappa * \lambda$ exists and has a continuous density f with respect to m. Then it follows from (5) that $k_\lambda = f$ locally m-almost everywhere. Since $k_\lambda - f$ is lower semi-continuous, the locally m-negligible set of points x where $k_\lambda(x) + - f(x) > 0$ is open and hence void (in view of the invariance of Haar measure). Having thus obtained $k_\lambda \leqslant f$ everywhere, we conclude that $\|\lambda\|^2 = k_\lambda(e) \leqslant f(e) < +\infty$, and hence $\lambda \in \mathcal{E}^+$, q.e.d. This result suggests a generalization of the theory of potentials with respect to a definite convolution kernel, in which the kernel function k is replaced by

(1) On any locally compact topological group the convolution $\omega = \lambda * \mu * \nu$ of three positive measures λ, μ, ν is defined by

$$\int \varphi \, d\omega = \iiint \varphi(xyz)\, d\lambda(x)\, d\mu(y)\, d\nu(z), \qquad \varphi \in \mathsf{C}_0^+,$$

provided the integral on the right is never infinite. It can be shown that the associative law

$$(\lambda * \mu) * \nu = \lambda * (\mu * \nu) = \lambda * \mu * \nu$$

holds in the sense that the existence of any one of these three expressions implies that of the others (provided $\lambda \neq 0$, $\nu \neq 0$).

(2) This result is similar to a result obtained by J. Deny [16], Théorème 3, for the group R^n. The conclusions are the same, but the type of definite kernel function considered by Deny differs *a priori* from that of the present paper. Subsequently, the quoted theorem of Deny implies that the kernel functions considered by him are likewise definite (in fact perfect) in our sense. Cf. Theorem 7.3 of the present paper.

a kernel measure $\varkappa$ which should be positive definite in a suitable sense. Cf., e.g., J., Deny [16], § 3.

7.2. *K-consistency of convolution kernels.* We denote by $\mathcal{U}^+$ the class of all measures $\mu \in \mathcal{M}^+(X)$ of compact support and of *unit* total mass: $\mu(X) = 1$. If we associate with every neighbourhood V of the unit element $e \in X$ the class $\mathcal{U}_V^+$ of all measures $\mu \in \mathcal{U}^+$ with $S(\mu) \subset V$, these classes $\mathcal{U}_V^+$ constitute the base of a filter Ψ on $\mathcal{M}^+$. This filter converges vaguely to ε (= the mass +1 placed at e). In fact, for every continuous function f on X,

$$\lim_{\mu \to \varepsilon} \int f \, d\mu = f(e) \quad \text{along } \Psi. \tag{1}$$

Similarly,

$$\lim_{\mu \to \varepsilon} \iint f(x, y) \, d\mu(x) \, d\mu(y) = f(e, e) \quad \text{along } \Psi \tag{2}$$

for every function $f \in C(X \times X)$. The verification, say of (2), is simple:

$$\left| \iint f(x, y) \, d\mu(x) \, d\mu(y) - f(e, e) \right| = \left| \iint (f(x, y) - f(e, e)) \, d\mu(x) \, d\mu(y) \right| \leqslant$$
$$\leqslant \max_{S(\mu) \times S(\mu)} |f(x, y) - f(e, e)| < \eta$$

provided $S(\mu)$ is contained in a sufficiently small neighbourhood $V = V_\eta$ of e.

LEMMA 7.2.1. *Given a K-definite kernel function k, let $\lambda \in \mathcal{E}$ denote an arbitrary measure of finite energy* (and of compact support unless $k \geqslant 0$). *Then $\lambda * \mu \in \mathcal{E}$, and*

$$\lim_{\mu \to \varepsilon} \| \lambda * \mu - \lambda \| = 0 \quad \text{along } \Psi.$$

Proof. It suffices to consider the case $\lambda \geqslant 0$. According to Lemma 6.4.1,

$$\| \lambda * \mu \|^2 = k_\lambda(\mu, \mu) = \iint k_\lambda(x y^{-1}) \, d\mu(x) \, d\mu(y),$$

and this is finite because k_λ is bounded and $\mu(X) = 1$. Similarly,

$$k(\lambda * \mu, \lambda) = k_\lambda(\mu, \varepsilon) = \int k_\lambda(x) \, d\mu(x).$$

Applying (2) and (1) with $f(x, y)$ and $f(x)$ replaced by $k_\lambda(x y^{-1})$ and $k_\lambda(x)$, respectively, we obtain along Ψ

$$\lim_{\mu} \| \lambda * \mu \|^2 = \lim_{\mu} k(\lambda * \mu, \lambda) = k_\lambda(e) = \| \lambda \|^2,$$

from which the strong convergence $\lambda * \mu \to \lambda$ along Ψ follows.

LEMMA 7.2.2. *If a K-definite kernel function k on the group X is Haar integrable over some neighbourhood of the identity $e \in X$, then k is locally Haar integrable; and every measure possessing a density of class $C_0(X)$ with respect to Haar measure has a continuous potential and a finite energy with respect to k. If k is not locally Haar integrable, then 0 is the only measure of finite energy.*

Proof. Suppose first k is Haar integrable over some neighbourhood W of e, and choose a neighbourhood V of e so that $VV^{-1} \subset W$. If $\varphi \in C_0^+(X)$ and $S(\varphi) \subset V$, then $\mu = \varphi m$ has the potential $k * \mu = k * \varphi$, the value of which at a point $x \in V$ is

$$\int k(xy^{-1})\,\varphi(y)\,dm(y) = \int_W \varphi(zx)\,k(z)\,dm(z).$$

The latter expression shows that $k * \mu$ is bounded and uniformly continuous in V. Integrating with respect to μ, we infer that $\mu \in \mathcal{E}^+$. Accepting for the moment the final assertion of the lemma, we conclude that k is, actually, locally Haar integrable in X, and hence we could take for V any relatively compact neighbourhood of the support of a given function $\varphi \in C_0^+(X)$. Thus we conclude that $\mu = \varphi m$ has a continuous potential and a finite energy. To finish the proof of the lemma, we suppose now that there exists a measure $\lambda \neq 0$ of finite energy. We may clearly assume that $\lambda \geq 0$, $\int d\lambda = 1$, and that $S(\lambda)$ is compact. For any given compact set $K \subset X$ we choose a function $\varphi \in C_0^+(X)$ so that $\varphi(x) = 1$ when $x \in S(\lambda)^{-1}K$. Then $\lambda * \varphi$ equals 1 everywhere in K. Writing $\mu = \varphi \check{m}$, we infer from (2), § 6.3, and the preceding lemma that

$$(\lambda * \varphi)\,\check{m} = \lambda * \mu \in \mathcal{E}^+.$$

The trace $\check{m}_K$ of the left invariant Haar measure $\check{m}$ on K coincides with the trace of $\lambda * \mu$ on K because $\lambda * \varphi = 1$ on K. Hence $\check{m}_K \in \mathcal{E}^+$, i.e., the kernel $k(xy^{-1})$ is integrable with respect to $\check{m} \otimes \check{m}$ over $K \times K$. Having thus shown that $k(xy^{-1})$ is locally integrable in $X \times X$, we finish by proving that the kernel function k itself is locally integrable in X. For any function $\varphi \in C_0^+(X)$, $\varphi \neq 0$, the function $\varphi(xy^{-1})\varphi(x)$ of (x, y) is of class $C_0^+(X \times X)$, and hence

$$\iint k(xy^{-1})\,\varphi(xy^{-1})\,\varphi(x)\,dm(x)\,dm(y) < +\infty,$$

i.e., $\int k\varphi\,d\check{m} \cdot \int \varphi\,dm < +\infty$, from which it follows that $\int k\varphi\,d\check{m} < +\infty$ because $\int \varphi\,dm \neq 0$.

The following application of Lemma 7.2.2 will not be used in the sequel, except for the observation that cap $X > 0$ if (and only if) the K-definite kernel function $k \geq 0$ is locally Haar integrable:

LEMMA 7.2.3. *Let k denote a locally Haar integrable, K-definite kernel function on the group X, and suppose that $k \geqslant 0$ if X is non-compact. Then*

(a) *The potential $k * \lambda$ is locally Haar integrable for every $\lambda \in \mathcal{E}$.*

(b) $\text{cap}_* N = 0$ *implies* $m_* (N) = 0$.

(c) $\text{cap}^* N = 0$ *implies* $m^* (N \cap K) = 0$ *for every compact set* $K \subset X$.

Proof. (a) The potential $k * \lambda$ of $\lambda \in \mathcal{E}$ is μ-integrable for every $\mu \in \mathcal{E}$, in particular for $\mu = \varphi m$, $\varphi \in C_0^+ (X)$. (b) For any $K \subset N$, the trace m_K of m upon K belongs to $\mathcal{E}^+$. Hence it follows from Lemma 2.3.1. that $m (K) = m_K (X) = 0$; and consequently $m_* (N) = 0$. (c) Let H denote a fixed, compact neighbourhood of the compact set K. Since $\text{cap}^* (N \cap K) = 0$, there are open sets $G \supset N \cap K$ with $\text{cap}\, G$ as small as we please. We may further assume that $G \subset H$. By definition,

$$\text{cap}\, G = 1/w (G) \geqslant \mu (X)^2 / \| \mu \|^2$$

for every non-zero measure $\mu \in \mathcal{E}^+$ concentrated on G. Taking $\mu = m_G$ (= the trace of m upon G), we obtain $m (G)^2 = m_G (X)^2 \leqslant \| m_G \|^2 \text{ cap } G$. If $k \geqslant 0$, $\| m_G \|^2 \leqslant \| m_H \|^2$ $(< + \infty)$, and hence $m (G)$ becomes as small as we please by suitable choice of G. This shows that $m^* (N \cap K) = 0$. The remaining case where X is compact, but k is of variable sign, is reduced in the usual way to the case $k \geqslant 0$. In fact, the condition $\text{cap}^* N = 0$ is not changed if k is replaced by $k + c$ for some constant c, as pointed out in § 2.3.

LEMMA 7.2.4. *Let k denote a K-definite kernel function on the group X. For any measure $\lambda \geqslant 0$ of finite energy* (and of compact support unless $k \geqslant 0$), *any neighbourhood W of the support of λ, and any number $\eta > 0$, there is a measure $\lambda' \geqslant 0$ possessing a continuous density of compact support contained in W, such that $\| \lambda' - \lambda \| < \eta$.*

Proof. In view of Lemma 7.2.2 we may suppose that k is locally Haar integrable. It follows easily from Lemma 2.2.2 and the remark thereto that there is a compact set $K \subset S (\lambda)$ such that $\| \lambda_K - \lambda \| < \eta/2$, where λ_K denotes the trace of λ upon K. It is easy to show that there is a compact neighbourhood V of the unit element $e \in X$ such that $KV \subset W$. According to Lemma 7.2.1 we may choose V so small that $\| \lambda_K * \mu - \lambda_K \| < \eta/2$ for every measure $\mu \in \mathfrak{U}_V^+$. Choose a function $\varphi \in C_0^+ (X)$ with $S (\varphi) \subset V$ and $\int \varphi \, d \check{m} = 1$. Writing $\mu = \varphi \check{m}$, we have $\mu \in \mathfrak{U}_V^+$. According to (2), § 6.3, the measure $\lambda' = \lambda_K * \mu$ has the density $\lambda_K * \varphi \in C_0^+ (X)$ with respect to $\check{m}$, and $S (\lambda') \subset W$. Clearly, $\| \lambda' - \lambda \| < \eta/2 + \eta/2$, and the proof is complete. Note that the potential $k * \lambda'$ of λ' is continuous by Lemma 7.2.2.

THEOREM 7.2. *Every K-definite convolution kernel* $k(xy^{-1})$ *on a locally compact topological group X is K-consistent. If X is compact, or if* $k \geqslant 0$ *and* $k(x) = 0$ *outside some compact subset of X, then* $k(xy^{-1})$ *is consistent.*

Proof. The latter part of the theorem follows from Lemma 3.4.2 in view of the preceding lemma with $W = X$, because the continuous potential $k * \lambda'$ has compact support when k and λ' both have compact supports. The former part follows from the corresponding criterion (b), § 5.3, for K-consistency, combined with the preceding lemma. In fact, for any compact set $K \subset X$ we may take $K_1 = W =$ an arbitrary compact neighbourhood of K.

COROLLARY. In case of a positive and K-definite convolution kernel, $\mathcal{E}_K^+$ is complete for every compact set $K \subset X$.

This follows from the criterion (a), § 5.3, for the completeness of $\mathcal{E}_K^+$. Disregarding the trivial case $k = 0$, we have $k(xx^{-1}) = k(e) > 0$, and hence the convolution kernel $k(xy^{-1})$ is strictly positive if at all positive. The question remains open whether this corollary holds in the case of K-definite convolution kernels of variable sign.

7.3. *Conditions for consistency of convolution kernels.* It remains to find conditions in order that a definite kernel function $k \geqslant 0$ on a locally compact, non-compact group X be *consistent* (and not merely K-consistent). Consistency is a global property, and it seems plausible that $k(x)$ should approach 0 in some sense as x approaches infinity in X.

LEMMA 7.3.1. *In case of a consistent convolution kernel* $(\geqslant 0)$ *on a locally compact group X, the capacity of any closed, non-compact subgroup of X is either* 0 *or* $+\infty$.

In particular, cap $X = +\infty$ is a necessary condition for consistency of a positive, locally Haar-integrable kernel function on a non-compact group X. If, for example, $X = R^n$, $n > 1$, this single condition is not sufficient.

Proof of Lemma 7.3.1. Replacing the kernel function by its restriction to the subgroup in question, one finds that it suffices to prove that cap $X = 0$ or $+\infty$. Let us suppose cap $X < +\infty$, and let Λ_X denote the non-void, vaguely compact class of all interior capacitary distributions on X (cf. Theorem 4.1). The invariance of the convolution kernel implies that Λ_X is invariant (as a whole) under right translations. (In the simple case where the kernel is strictly definite, the unique interior capacitary distribution λ on X is therefore a right invariant measure on X; and since Haar measure on a non-compact group is infinite, whereas $\lambda(X) =$ cap $X < +\infty$, we conclude that $\lambda = 0$; that is, cap $X = 0$.) In the general case, let λ denote a fixed interior capacitary distribution on X. The mapping $x \to \lambda * \varepsilon_x$ carries the filter of "neighbourhoods of infinity" in X (= the complements of all relatively compact subsets of X) into the base of a filter Φ on the vaguely compact set Λ_X. Let μ denote any vague cluster point for Φ. Then $\mu \in \Lambda_X$, and hence $\mu(X) = \text{cap } X < +\infty$. We complete the proof by showing that the translates $\lambda * \varepsilon_x$ converge vaguely to 0 as $x \to \omega$ (the Alexandrov point at infinity adjoined to X); and hence $\mu = 0$. Corresponding to a given number $\eta > 0$ we determine a compact set $K \subset X$ so that $\lambda(\complement K) < \eta$. This is possible because $\lambda(X) = \text{cap } X < +\infty$ by assumption. For any function $\varphi \in C_0^+(X)$, say $\varphi \leqslant 1$, the set $K^{-1} S$ is compact, S being the support of φ. Since X is non-compact, there are points $x \in \complement(K^{-1} S)$. For any such point x, $\varphi(yx)$ vanishes for $y \in K$, and hence

$$\int \varphi \, d(\lambda * \varepsilon_x) = \int \varphi(yx) \, d\lambda(y) \leqslant 1 \cdot \lambda(\complement K) < \eta.$$

Letting $x \to \omega$, we obtain $\int \varphi \, d\mu \leqslant \eta$; and hence $\int \varphi \, d\mu = 0$, i.e., $\mu = 0$.

The condition $\text{cap } X = +\infty$ amounts to the requirement that a suitable mean-value of the kernel k should equal 0. As an illustration we state without proof the following lemma, in which $X = R^n$ (considered as a group under addition):

LEMMA 7.3.2. *If $k \geqslant 0$ is a definite kernel function on R^n, then*

$$\frac{1}{\text{cap }(R^n)} = \lim_{t \to \infty} \frac{1}{t^n} \int_{-t}^{t} \cdots \int_{-t}^{t} \prod_{i=1}^{n} \left(1 - \frac{|x_i|}{t}\right) k(x) \; dx_1 \ldots dx_n.$$

If $M(r)$ denotes the mean-value of $k(x)$ over the sphere $|x| = r$, the condition $\text{cap } X = +\infty$ *is equivalent to the limit relation*

$$\lim_{r \to \infty} \frac{1}{r^n} \int_0^r \left(1 - \frac{\varrho}{r}\right) M(\varrho) \, \varrho^{n-1} \, d\varrho = 0.$$

In the way of *sufficient* conditions for a convolution kernel to be consistent, one has the following theorem due to J. Deny [16], and, moreover, a result described in the next section (Theorem 7.4).

THEOREM 7.3. *A regular kernel function $k \geqslant 0$ on R^n satisfying Condition (A) in* Deny [16], § 1, *is perfect.*

Proof. The only difficulty is to pass from the positive character implied by Cond. (A) to that of the present paper. In the case of a *regular* kernel function this is possible by virtue of Théorème 3 in Deny [16], p. 97. In fact, to prove that k is definite, we must show that

$$k(\mu^+, \mu^+) + k(\mu^-, \mu^-) \geq 2\,k(\mu^+, \mu^-)$$

for every measure μ. It suffices, of course, to consider measures μ such that $\mu^+, \mu^- \in \mathcal{E}^+$. According to the quoted theorem of Deny, the class $\mathcal{E}^+$ coincides with the class $\mathcal{E}'$ of all positive measures λ such that the convolution $\check{\lambda} * (k\,m) * \lambda$ exists and possesses a continuous density with respect to Lebesgue measure m in R^n. Moreover, the mutual energy $k(\lambda, \nu)$ of any two measures $\lambda, \nu \in \mathcal{E}^+ = \mathcal{E}'$ is the value (λ, ν) at the origin of the continuous density of $\check{\lambda} * (k\,m) * \nu$ (Deny [16], formula (2), p. 98. See also Théorème 2, 2°, p. 93 f.). Applying this with λ and ν replaced by μ^+ or μ^-, one obtains

$$k(\mu, \mu) = (\mu, \mu) \geq 0,$$

with equality only for $\mu = 0$. Consequently, k is strictly definite in the sense of the present paper. It follows now from Théorème 2 of the quoted paper that $\mathcal{E}^+$ is complete, and from the last note on p. 94 that strong convergence in $\mathcal{E}^+$ implies vague convergence. Thus k is perfect. (Alternatively, one might establish the consistency of k by application of Lemma 3.4.2 in the present paper.) Cond. (A) is not necessary for a kernel to be perfect (cf. § 8.2).

7.4. *The case* $k = \check{h} * h$. Let $h \geq 0$ denote a lower semi-continuous function on a locally compact group X, and put $k = \check{h} * h$ (cf. (3), § 6.3). Then the convolution kernel

$$k(x y^{-1}) = \int h(t x^{-1})\, h(t y^{-1})\, d m(t)$$

is obtained by superposition (in the sense of § 3.5) of the kernels

$$k_t(x, y) = h(t x^{-1})\, h(t y^{-1}), \qquad t \in X,$$

each of which is obviously definite, the energy being

$$k_t(\mu, \mu) = \left\{\int h(t x^{-1})\, d\mu(x)\right\}^2$$

(= the value of $(h * \mu)^2$ at the point t). Note that k_t is in general inconsistent (cf. Ex. 3, § 8.3). It follows from § 3.5 that k is *definite*, the corresponding energy being

$$\|\mu\|^2 = k(\mu, \mu) = \int (h * \mu)^2\, d m. \tag{1}$$

Since $\mu \in \mathcal{E}$ is equivalent to $|\mu| \in \mathcal{E}^+$, that is $k(|\mu|, |\mu|) < \infty$, we infer from (1) that the class $\mathcal{E}$ of all measures μ of finite energy with respect to k is determined by the condition $h*|\mu| \in \mathfrak{L}^2(m)$. (1)

If $k(e)=0$, then $\int h^2\,dm=0$, and hence $h=0$ because h is lower semi-continuous. (The open set $\{x \in X : h(x)>0\}$ is m-negligible and hence void.) From Lemma 7.2.2 follows that k is locally Haar integrable if and only if $h*\psi\ (=h*(\psi m))$ is in $\mathfrak{L}^2(m)$ for every $\psi \in C_0^+(X)$.

In the special case where h (and hence k) vanishes outside some compact subset of X, we infer from Theorem 7.2 that k is consistent. In the general case, the same conclusion is possible at least if X is *Abelian*:

THEOREM 7.4. *Let X denote a locally compact Abelian group, and $h \geqslant 0$ a lower semi-continuous function on X. The kernel function $k=\check{h}*h$ is consistent, and $\mathcal{E}^+$ is complete.*

Proof. We may, of course, suppose that k (and hence h) does not vanish identically. Thus the convolution kernel $k(xy^{-1})$ is strictly positive: $k(e)>0$, and we infer from Lemma 3.3.1 that $\mathcal{E}^+$ is complete if k is consistent. In the proof of the consistency, we may assume that k is locally Haar integrable, that is, as pointed out above,

$$h*\psi \in \mathfrak{L}^2(m) \qquad \text{for every } \psi \in C_0^+(X). \tag{2}$$

Let Φ denote a strong Cauchy filter on $\mathcal{E}^+$, or just as well on the part of $\mathcal{E}^+$ determined by $\|\mu\|^2 \leqslant M$ for some constant M; and suppose Φ converges vaguely to some measure μ_0 (cf. Condition (C′), § 3.3). According to (1), the mapping $\mu \to h*\mu$ carries $\mathcal{E}$ isometrically into $\mathfrak{L}^2(m)$. The image of Φ by this mapping is, therefore, the base of a Cauchy filter on $\mathfrak{L}^2(m)$, and converges to some function $f \in \mathfrak{L}^2(m)$ on account of the Riesz-Fischer theorem (cf. Bourbaki [4], Chap. IV, § 3, th. 2):

$$\lim_{\mu} \int (h*\mu - f)^2\,dm = 0 \qquad \text{along } \Phi. \tag{3}$$

Our task is to prove that $h*\mu_0=f$ m-almost everywhere. Denoting throughout the rest of the proof by φ an arbitrary function of class C_0^+ such that $\varphi \leqslant h$, we begin by showing that $\varphi*\mu \to \varphi*\mu_0$ *weakly* in $\mathfrak{L}^2(m)$, that is,

(1) It should be observed that the weaker condition $h*\mu \in \mathfrak{L}^2$ is necessary, but not sufficient for a measure μ of variable sign to be of finite energy with respect to $k=\check{h}*h$. In fact, the energy of such a measure μ does not necessarily exist in the sense of the definition (1), p. 149. Cf. the end of note (1) on p. 206; or the final observation in Example 9, p. 213.

$$\int (\varphi * \mu)\, v\, dm \to \int (\varphi * \mu_0)\, v\, dm \quad \text{along } \Phi \tag{4}$$

for every function $v \in \mathfrak{L}^2(m)$. Since

$$\int (\varphi * \mu)^2\, dm \leq \int (h * \mu)^2\, dm = \|\mu\|^2 \leq M$$

for all relevant measures μ, it suffices to verify (4) for functions $v \in C_0$, the class C_0 being everywhere dense in $\mathfrak{L}^2$. According to (4), (5), and (6), § 6.3,

$$\int (\varphi * \mu)\, v\, dm = (v * \varphi * \mu)(e) = \int (\check{\varphi} * v)\, d\mu.$$

Since $\check{\varphi} * v \in C_0$, (4) now follows from the vague convergence $\Phi \to \mu_0$.

According to (3), $h * \mu \to f$ strongly and hence weakly in $\mathfrak{L}^2$:

$$\int (h * \mu)\, g\, dm \to \int f g\, dm \quad \text{along } \Phi \tag{5}$$

for every $g \in \mathfrak{L}^2(m)$. Taking

$$g = \check{\varphi} * \psi, \qquad (\varphi, \psi \in C_0^+, \varphi \leq h) \tag{6}$$

we obtain, since $\varphi * h = h * \varphi$ when X is Abelian,

$$\int (h * \mu)(\check{\varphi} * \psi)\, dm = (\check{\psi} * \varphi * h * \mu)(e) = \int (\varphi * \mu)(\check{h} * \psi)\, dm. \tag{7}$$

A similar computation shows that

$$\int (\check{h} * \psi)^2\, dm = \int (h * \psi)^2\, dm$$

because $h * \check{h} = \check{h} * h$. In view of (2), this implies $\check{h} * \psi \in \mathfrak{L}^2(m)$. It follows now from (6), (7), and (4) that, along Φ,

$$\int (h * \mu)\, g\, dm \to \int (\varphi * \mu_0)\, v\, dm = \int (h * \mu_0)\, g\, dm$$

for $v = \check{h} * \psi$. Combining this with (5), we obtain

$$\int (h * \mu_0)\, g\, dm = \int f g\, dm \tag{8}$$

for every function $g = \check{\varphi} * \psi$ of the type (6). The class $\mathcal{G}$ of all such functions g is contained in $C_0^+(X)$ and is invariant under right translations. Moreover, $\mathcal{G}$ contains non-zero functions of arbitrarily small support. Hence we infer from a result due to

H. Cartan that $\mathcal{G}$ is *total* in C_0^+. (1) It follows, therefore, from (8) that the measures $(h*\mu_0)\cdot m$ and $f\cdot m$ are identical, and hence their densities $h*\mu_0$ and f coincide locally m-almost everywhere in X. Since the two densities are of class $\mathcal{L}^2(m)$, we conclude that, actually, $h*\mu_0=f$ m-almost everywhere. This completes the proof.

Remark. The above theorem may be extended slightly by replacing Haar measure m by some measure $\tau=q\cdot m$ possessing a density $q=q(x)>0$ which is continuous and multiplicative:

$$q(xy)=q(x)\,q(y).$$

For any lower semi-continuous function $g\geqslant 0$ on X, the kernel

$$k(x,y)=\int g(xt)\,g(yt)\,d\tau(t) \tag{9}$$

is then definite according to § 3.5. Moreover, k is *consistent* if X is Abelian. Writing $h=q^{-1/2}\check{g}$, one obtains, in fact,

$$k(x,y)=q(x)^{-1/2}\,q(y)^{-1/2}\,(\check{h}*h)(xy^{-1}),$$

and the consistency of k follows from that of $\check{h}*h$ as explained in § 5.1. As an illustration we mention the following kernel on the multiplicative group of real numbers >0:

$$k(x,y)=(x+y)^{-\alpha}=\frac{1}{\Gamma(\alpha)}\int_0^\infty \exp(-tx)\exp(-ty)\,t^{\alpha-1}\,dt.$$

Here $\alpha>0$ is a constant, $q(t)=t^\alpha$, and $dm=t^{-1}dt$. We show in Ex. 7, § 8.3, that this kernel is strictly definite, and hence *perfect*.

8. Examples

8.1. *Kernels of order* α. Let X denote a locally compact topological group, and let h_α denote a symmetric, lower semi-continuous, locally Haar integrable function $\geqslant 0$ on X, depending on a parameter $\alpha\,(0<\alpha<A)$ in such a way that

$$h_\alpha * h_\beta = h_{\alpha+\beta}, \qquad (\alpha>0,\ \beta>0,\ \alpha+\beta<A) \tag{1}$$

and further that the measures $h_\alpha m$ vary continuously with α in the vague topology, and $h_\alpha m\rightarrow\varepsilon$ vaguely as $\alpha\rightarrow 0$. This type of a family of kernel functions was studied

(1) Cf. H. Cartan [8]; or [9], p. 78. For any total subclass $\mathcal{G}\subset C_0^+(X)$, the vector space $\mathcal{V}$ consisting of all finite linear combinations of functions from $\mathcal{G}$ is *positive rich* in the sense of Bourbaki [4], Chap. III, § 2, N° 5, and hence the identity stated above in the text follows from Bourbaki [4], Chap. III, prop. 2.

by H. Cartan [9], who proved (§ IV) that each h_α is *K-perfect*. If the group X is Abelian, it follows from Theorem 7.4 of the present paper that each h_α is *consistent* and that $\mathcal{E}^+$ is complete. If it can be shown that the kernel functions h_α are strictly definite, we conclude from Theorem 3.3 that they are *perfect*. This is the case, e.g., by the kernels of order α of M. Riesz:

$$h_\alpha(x)=\frac{1}{H_n(\alpha)}|x|^{\alpha-n}, \qquad (x\in R^n,\ 0<\alpha<n), \tag{2}$$

where
$$H_n(\alpha)=2^\alpha\,\pi^{\frac{1}{2}n}\,\frac{\Gamma(\frac{1}{2}\alpha)}{\Gamma(\frac{1}{2}(n-\alpha))} \tag{3}$$

(cf. M. Riesz [26], in particular p. 10 f). His proof of the strict definiteness of these kernel functions h_α is based on the composition formula (1) together with the following identity, in which Δ denotes the Laplace operator:

$$-\Delta\, h_\alpha(x)=h_{\alpha-2}(x) \qquad (x\neq 0,\ \alpha>2). \tag{4}$$

Using these tools, Riesz showed that every sufficiently differentiable function of compact support in R^n has the form $h_\alpha * \mu$ for a suitable measure μ of finite energy with respect to h_α, α being given. Since smooth functions of compact support form a rich subclass of C_0, it follows immediately that h_α is strictly definite (cf. Lemma 3.4.3). Occasionally it is convenient to consider these kernels of order α for values $\alpha\geqslant n$. For $n<\alpha<n+2$, the expression (3) is well-defined, and h_α is continuous throughout R^n. For $\alpha=n$ one is led to the logarithmic kernel in R^n by the following definition (cf. M. Riesz, loc. cit.)

$$h_n(x)=\lim_{\varepsilon\to 0}\tfrac{1}{2}\,(h_{n-\varepsilon}(x)+h_{n+\varepsilon}(x))=\frac{\omega_n}{(2\pi)^n}\log\frac{1}{|x|}.$$

Here ω_n denotes the surface of the unit sphere in R^n. The composition formula (1) does not subsist for $\alpha+\beta\geqslant n$. Frostman's maximum principle is fulfilled for $\alpha\leqslant 2$ (cf. Frostman [18], p. 68). Apart from the constant factor, the *Newtonian kernel* corresponds to the kernel function h_2:

$$h_2(x)=\begin{cases}\dfrac{1}{(n-2)\,\omega_n}|x|^{2-n} & (n\neq 2)\\[2ex] \dfrac{1}{2\pi}\log\dfrac{1}{|x|} & (n=2).\end{cases}$$

A related family of perfect kernels "of order α" in R^n, $0<\alpha<+\infty$, is determined by the kernel functions

$$g_\alpha(x) = \frac{1}{G_n(\alpha)} \frac{K_{\frac{1}{2}(n-\alpha)}(r)}{r^{\frac{1}{2}(n-\alpha)}}, \quad (r = |x|),$$

where
$$G_n(\alpha) = (2\pi)^{\frac{1}{2}n} 2^{\frac{1}{2}\alpha - 1} \Gamma(\tfrac{1}{2}\alpha)$$

and $K_\lambda(r)$ is the modified Bessel function of the third kind of order λ. Again (1) is fulfilled (now for all $\alpha > 0$, $\beta > 0$). In analogy with (4),

$$(1-\Delta) g_\alpha = g_{\alpha-2} \qquad (x \neq 0, \alpha > 2).$$

A theory of potentials with respect to these kernels has recently been developed by N. Aronszajn and K. T. Smith. Both of these examples are better understood in the light of the theory of distributions (cf. Schwartz [27], Chap. II, § 3, Ex. 2).

8.2. *Further types of kernels on R^n depending only on $|x-y|$.* We begin by considering the following simple kernel function k_t on R^1, $t > 0$ being given:

$$k_t(x) = (t - |x|)^+ = \max\{t - |x|, 0\}.$$

If we denote by j the characteristic function associated with the open interval $J = \{x \in R^1 : 0 < x < t\}$, we find $k_t = \check{j} \ast j$, and hence k_t is consistent (Theorem 7.2 or 7.4). It is easily verified that k_t is strictly definite, (1) and hence *perfect*. Note that

(1) Suppose $\mu \in \mathcal{E}$ (that is, $j \ast |\mu| \in \mathfrak{L}^2$) and $k_t(\mu, \mu) = 0$. Then $j \ast \mu = 0$ almost everywhere. For any function $\varphi \in \mathfrak{C}_0^+$ the continuous function $f = \varphi \ast \mu$ fulfills the condition

$$j \ast f = \varphi \ast (j \ast \mu) = 0 \text{ everywhere.}$$

In view of the definition of j, this means that the primitive $\int_0^x f(y)\,dy$ has the period t:

$$\int_{x-t}^{x} f(y)\,dy = (j \ast f)(x) = 0 \qquad \text{for every } x.$$

It follows that f, and hence $|f|$, is periodic with the period t, and $j \ast |f|$ is therefore a constant. Since

$$j \ast |f| \leq \varphi \ast (j \ast |\mu|) \in \mathfrak{L}^2,$$

this constant must equal 0, so that $|f| = 0$. Consequently,

$$\int \check{\varphi}\, d\mu = (\varphi \ast \mu)(0) = f(0) = 0$$

for every $\varphi \in \mathfrak{C}_0^+$, and we conclude that $\mu = 0$.

Observe that the measure ν of density $\sin(2\pi x/t)$ with respect to Lebesgue measure fulfills the condition $k \ast \nu = j \ast \nu = 0$; and hence, formally, $\|\nu\|^2 = 0$. The energy $\|\nu\|^2$ of ν is, however, not defined because $j \ast |\nu|$ is not square integrable over the real line.

Theorem 7.3 is inapplicable here because the Fourier transform of k_t is the Fejér kernel, which has zeros.

More generally, let $k \geqslant 0$ denote an even function on R^1 such that $k(x)$ is continuous and *convex* for $0 < x < +\infty$, and further

$$\lim_{|x|\to\infty} k(x) = 0; \qquad \lim_{|x|\to 0} k(x) = k(0) \leqslant +\infty.$$

Clearly, $k(x)$ is decreasing for $x \geqslant 0$, and hence $p(x) = D_+ k(x) \leqslant 0$ for $x > 0$. The convexity of k implies that p is increasing. By partial integration we obtain

$$k(x) = k(|x|) = -\int_{|x|}^{\infty} p(t)\,dt = \int_{|x|}^{\infty} (t - |x|)\,dp(t) = \int_{0}^{\infty} k_t(x)\,dp(t)$$

because $p(t) = o(t^{-1})$ as $t \to +\infty$. (1) Thus k can be obtained by superposition of the perfect kernels k_t, and we conclude from § 3.5 that k is perfect (the only exception being $k = 0$).

Next, one may study similar kernel functions on R^n, $n \geqslant 2$. Denoting be h the Newtonian kernel function on R^n, we now define

$$k_t(x) = (h(x) - t)^+ = \max\{h(x) - t, 0\},$$

where $t \geqslant 0$ for $n \geqslant 3$, whereas t is arbitrary in the case $n = 2$. It is possible to prove directly by elementary methods that these kernel functions k_t are strictly K-definite, and since they have compact supports (except for $t = 0$, $n \geqslant 3$, where $k_t = h$), we conclude from Theorem 7.2 that they are consistent (and K-perfect). Actually, each k_t is *perfect* by virtue of Theorem 7.3 (which was inapplicable for $n = 1$). An explicit calculation of the Fourier transform of k_t shows, in fact, that the regular kernel function k_t fulfills Condition (A) in Deny [16], § 1.

Finally, one may consider a kernel function $k \geqslant 0$ on R^n, $n \geqslant 2$, depending only on $|x| = r$ and in such a way that $k(x)$ is continuous and subharmonic for $x \neq 0$, $k(0) = +\infty$, and $k(x) \to 0$ as $|x| \to \infty$. (The subharmonicity means that k should be a convex function of the Newtonian kernel h.) This type of kernel function was studied by K. Kunugui [20] and N. Ninomiya [22], who proved that these kernels fulfill Frostman's maximum principle and are strictly K-definite. If we observe that each of these kernels can be obtained by superposition in the sense of § 3.5 of the above special kernels $k_t = (h - t)^+$, which are perfect, we conclude from § 3.5 that the kernels of Kunugui are *perfect*.

(1) In fact,

$$t\,|p(t)| = 2\,(t - \tfrac{1}{2}t)\,|p(t)| \leqslant 2\,k(\tfrac{1}{2}t) - 2\,k(t) \to 0.$$

For $n \geqslant 3$, the above class of kernel functions is a proper subclass of the class considered by T. Ugaheri [28], viz. the class of all functions of the form

$$k(x) = \varphi(r) \cdot h(x) = \varphi(r) \cdot r^{2-n},$$

where φ denotes an arbitrary decreasing function of $|x| = r$ (continuous from the right). As shown by Ugaheri, these kernel functions are strictly K-definite (but they do not all fulfill Frostman's maximum principle). Actually, the kernels $k \not\equiv 0$ of Ugaheri are *perfect*. This may be shown in the manner indicated above for the Kunugui kernels. In fact, k can be obtained by superposition of special Ugaheri kernels k_t obtained by taking $\varphi(r) = 1$ for $r < t$, $\varphi(r) = 0$ for $r \geqslant t$ $(0 < t \leqslant +\infty)$.

8.3. *Miscellaneous examples.* We bring a number of examples designed to illustrate various points of the preceding theory.

Example 1. The simplest example of a *definite,* but *inconsistent* kernel on a locally compact, non-compact space X is the constant $k(x, y) = 1$. Clearly, $k(\mu, \mu) = \mu(X)^2 \geqslant 0$, and $\mathcal{E}$ consists of all bounded measures. To see that $k = 1$ is inconsistent, we observe that the measure ε_x (= the mass $+1$ placed at x) converges vaguely to 0 as x approaches the Alexandrov point ω at infinity adjoined to X. Since the measures ε_x all belong to one and the same equivalence class in $\mathcal{E}^+$, the mapping $x \to \varepsilon_x$ carries the filter of neighbourhoods of ω in X into the base of a strong Cauchy filter Φ on $\mathcal{E}^+$. If $k = 1$ were consistent, Φ should converge strongly to its vague limit 0, but this is not the case since $\|\varepsilon_x\| = 1$. Nevertheless, $\mathcal{E}^+$ is easily shown to be strongly complete. Note that the kernel 1 is a convolution kernel if X is a group.

Example 2. The kernel

$$\cos(x - y) = \cos x \cos y + \sin x \sin y$$

on the additive group R^1 is an example of a *K-definite convolution kernel* which is *not strictly pseudo-positive.* (The only such *positive* kernel is 0.) In fact, the positive measure $\varepsilon_0 + \varepsilon_\pi$ has the energy 0. The compact set consisting of the two points 0 and π has infinite capacity, but each of the two points forms a set of capacity 1. The space $\mathcal{E}_K^+$ is strongly complete for every compact set K.

Example 3. As a generalization of Ex. 1 we consider kernels of the form $f \otimes f$ on a locally compact space X:

$$k(x, y) = f(x) f(y).$$

We suppose f is lower semi-continuous and $0 \leqslant f(x) \leqslant +\infty$. Such a kernel is positive and definite. It is strictly pseudo-positive if $f(x) > 0$ for all x. It is not strictly definite (unless X reduces to a single point). $\mathcal{E}$ consists of all measures μ such that $\int f d|\mu| < +\infty$. The energy and potential of $\mu \in \mathcal{E}$ are

$$\|\mu\|^2 = \left(\int f d\mu\right)^2, \qquad k(x, \mu) = f(x) \cdot \int f d\mu.$$

$\mathcal{E}^+$ is always complete, but k is consistent if and only if $f \in C_0^+(X)$. (The sufficiency of this condition follows from Lemma 3.4.2. The necessity can be proved by considerations similar to those employed in Ex. 1: With every point x of the open set $G = \{x \in X : f(x) > 0\}$ one associates the measure $\mu_x = \varepsilon_x / f(x)$.) In any case, it is easily shown that all sets are capacitable, and

$$\operatorname{cap}_* A = \operatorname{cap}^* A = \sup_{x \in A} \{1/f(x)^2\}.$$

Example 4. A simple example of a *strictly definite convolution kernel* $k \geqslant 0$ which is *not perfect* is obtained by adding the definite kernel 1 (cf. Ex. 1) to the Newtonian kernel h in R^3:

$$k(x, y) = |x - y|^{-1} + 1.$$

Since $k(\mu, \mu) \geqslant h(\mu, \mu)$ for any measure μ of finite energy with respect to k, the kernel k has the property (P_2), § 3.3 (because h has this property). We proceed to show that k does not have the property (P_1) of completeness of $\mathcal{E}^+$. Let σ_r denote the uniform distribution of unit mass on the sphere $\{x \in R^3 : |x| = r\}$. Then the energy of $\sigma_r - \sigma_s$ is the same whether taken with respect to the kernel k or to the Newtonian kernel h because the contribution from the constant 1 is $\left\{\int d(\sigma_r - \sigma_s)\right\}^2 = 0$. Using a classical property of the Newtonian potential of σ_r, we obtain

$$\|\sigma_r - \sigma_s\|^2 = \frac{1}{r} - \frac{1}{s} \qquad \text{for } s > r,$$

and $\{\sigma_n\}_1^\infty$ is therefore a strong Cauchy sequence with respect to both kernels. In view of (P_2), the only possible strong limit of $\{\sigma_n\}$ is the vague limit 0. However, $k(\sigma_n, \sigma_n) = 1 + n^{-1}$ does not approach 0. Consequently, $\mathcal{E}^+$ is incomplete.

Denoting the interior capacities associated with the kernels h and k by γ_* and cap_*, respectively, one has

$$\operatorname{cap}_* A = \frac{\gamma_*(A)}{1 + \gamma_*(A)},$$

and similarly for the two exterior capacities. In particular, cap $X = 1$, and we have thus obtained an alternative proof of the inconsistency of k (cf. Lemma 7.3.1 or Lemma 7.3.2). There is no interior capacitary distribution associated with the entire space $X = R^3$. Nevertheless, the capacitable sets are the same for the two kernels h and k.

Example 5. As an example of a *strictly definite* kernel $k \geqslant 0$ which is *not K-perfect,* we consider the kernel

$$k(x, y) = |x - y|^{-1} + b(x)\, b(y)$$

on $X = R^3$. Here b denotes the characteristic function associated with the open unit ball $B = \{x \in R^3\colon |x| < 1\}$. The energy $k(\mu, \mu)$ of a measure μ is obtained from the Newtonian energy $h(\mu, \mu)$ by adding $\mu(B)^2$. The space $\mathcal{E}^+$ consists of the same measures whether taken with respect to h or to k. Since $h(\mu, \mu) \leqslant k(\mu, \mu)$, k fulfills (P_2), § 3.3. However, $\mathcal{E}_B^+$ is incomplete. In fact, when $r \to 1$ through some sequence of numbers $r < 1$, the sequence $\{\sigma_r\}$ (cf. Ex. 4) is a strong Cauchy sequence in $\mathcal{E}_B^+$ which converges vaguely, but not strongly to σ_1. Observe also that the open set B is of finite capacity $\frac{1}{2}$, but there is no interior capacitary distribution associated with B.

Example 6. The continuous kernel $k(x, y) = xy/(2 - xy)$ on the compact interval $X = \{x \in R^1\colon 0 \leqslant x \leqslant 1\}$ is positive, but not strictly positive because $k(0, 0) = 0$. The identity

$$k(x, y) = \sum_{p=1}^{\infty} 2^{-p} x^p y^p \qquad (0 \leqslant x, y \leqslant 1)$$

shows that k is definite and, in fact, *consistent* because each function $x^p y^p$ is a consistent kernel (cf. §§ 3.4 and 3.5). Note that $0 \leqslant k(x, y) \leqslant xy$. We proceed to prove that $\mathcal{E}^+$ is *incomplete*. Let $0 < a_i \leqslant 1$; $a_{i+1} < a_i$; and $\sum_{i=1}^{\infty} a_i < +\infty$. The measures

$$\mu_n = \varepsilon_{a_1} + \varepsilon_{a_2} + \cdots + \varepsilon_{a_n}$$

form a strong Cauchy sequence in $\mathcal{E}^+$ because

$$\sum_{i=1}^{\infty} \|\varepsilon_{a_i}\| = \sum_{i=1}^{\infty} k(a_i, a_i)^{\frac{1}{2}} \leqslant \sum_{i=1}^{\infty} a_i < +\infty.$$

If $\mu \in \mathcal{E}^+$ denotes some strong limit of $\{\mu_n\}$, then, for every $p = 1, 2, \ldots$,

$$2^{-p}\left\{\int_0^1 x^p\, d(\mu_n - \mu)\right\}^2 \leqslant \|\mu_n - \mu\|^2 \to 0 \qquad \text{as } n \to \infty.$$

Introducing the measures $\lambda_n = x \cdot \mu_n = \sum_{i=1}^{n} a_i \varepsilon_{a_i}$ and $\lambda = x \cdot \mu$, we obtain

$$\lim_{n\to\infty} \int_0^1 x^{p-1}\, d\lambda_n = \int_0^1 x^{p-1}\, d\lambda, \qquad p = 1, 2, \ldots,$$

and hence

$$\lim_n \int_0^1 P(x)\, d\lambda_n(x) = \int_0^1 P(x)\, d\lambda(x) \tag{1}$$

for every polynomial P. The inequality

$$2^{-1}\left\{\int_0^1 d\lambda_n\right\}^2 = 2^{-1}\left\{\int_0^1 x\, d\mu_n\right\}^2 \leqslant \|\mu_n\|^2$$

shows that the measures λ_n are uniformly bounded, and hence it follows from (1) that $\lambda_n \to \lambda$ vaguely; that is,

$$\lambda = \sum_{n=1}^{\infty} a_n \varepsilon_{a_n}.$$

If A_n denotes the set $\{a_1, a_2, \ldots, a_n\}$, we obtain

$$\mu(A_n) = \int_{A_n} x^{-1}\, d\lambda = n,$$

which contradicts the finiteness of $\mu(X)$ on a compact space X.

Example 7. For any number $\alpha > 0$ the kernel

$$k(x, y) = (x+y)^{-\alpha} = \frac{1}{\Gamma(\alpha)} \int_0^{\infty} e^{-t(x+y)}\, t^{\alpha-1}\, dt$$

on the semi-axis $X = \{x \in R^1\colon\ 0 < x < +\infty\}$ is *consistent* (cf. the remark following Theorem 7.4). The energy of a measure μ is

$$\|\mu\|^2 = \frac{1}{\Gamma(\alpha)} \int_0^{\infty} (L\mu)^2\, t^{\alpha-1}\, dt,$$

where $L\mu$ denotes the Laplace transform of μ:

$$(L\mu)(t) = \int_0^{\infty} e^{-tx}\, d\mu(x), \qquad 0 < t < +\infty.$$

The space $\mathcal{E}$ consists of all measures μ for which $L|\mu|$ is square integrable with respect to $t^{\alpha-1}\,dt$. For such a measure μ, $L\mu$ converges absolutely for $t>0$ and represents a continuous (in fact, analytic) function of t. If $\|\mu\|^2=0$, then $L\mu=0$, and hence $\mu=0$ (cf., e.g., D. V. Widder [32], § 6). Consequently, k is *perfect.*

For a positive measure μ, $k(x,\mu)$ decreases and $x^\alpha k(x,\mu)$ increases with $x>0$. Hence $k(x,\mu)$ is finite for all $x>0$, unless $k(x,\mu)\equiv+\infty$. This latter possibility cannot occur if $\mu\in\mathcal{E}^+$. We have thus shown that $(x+y)^{-\alpha}$ is a μ-integrable function of y for fixed $x>0$, provided $\mu\in\mathcal{E}^+$, or equally well: $\mu\in\mathcal{E}$. *A fortiori*, the same is true of the function $(x+y)^{-\alpha-1}$, and hence there is no difficulty in verifying that $k(x,\mu)$ is analytic for $Re\,(x)>0$ provided $\mu\in\mathcal{E}$. In particular, $k(x,\mu)$ can only be of class $C_0(X)$ if $k(x,\mu)\equiv 0$; that is, if $\mu=0$. This shows that the consistency condition formulated in Lemma 3.4.2 is not a necessary one. In the next example we exhibit a related *convolution kernel* with similar properties.

Example 8. For any number $\alpha>0$ the kernel function

$$k(x)=\left(2\cosh\frac{x}{2}\right)^{-\alpha}$$

on the additive group R^1 of real numbers may be represented as $k=\check{h}*h$, where

$$h(x)=(\Gamma(\alpha))^{-\frac{1}{2}}\exp\left(\tfrac{1}{2}\alpha x-e^x\right).$$

The corresponding convolution kernel $k(x-y)$ is therefore consistent (Theorem 7.4). By the substitution $x=\log u$, $y=\log v$, $k(x-y)$ is transformed into the kernel

$$\frac{(uv)^{\frac{1}{2}\alpha}}{(u+v)^\alpha}$$

on the positive semi-axis $0<u,\ v<+\infty$. In view of an observation in § 5.1, this new kernel is perfect because $(u+v)^{-\alpha}$ is a perfect kernel (cf. the preceding example). Consequently, the original kernel function k is *perfect*, too. For every $\mu\in\mathcal{E}$ the potential $k*\mu$ is an entire analytic function (and hence never of class C_0 unless $\mu=0$).

Example 9. In order to show that the condition for strict definiteness formulated in Lemma 3.4.3 is not a necessary one, we consider the following kernel function $k=k(n)$ on the additive group N of integers with the discrete topology: $k(n)=2$ for $n=0$; $k(n)=1$ for $n=\pm1$; and $k(n)=0$ elsewhere. (This kernel function is a discrete analogue of the kernel function k_t studied in § 8.2 in the case $n=1$.) If j denotes the characteristic function associated with the set consisting of the points $n=0$ and

$n=1$, we have $k=\check{j}*j$. Hence k is consistent, and the energy of a measure μ on N is

$$\|\mu\|^2=\sum_n (\mu_n+\mu_{n+1})^2.$$

(We denote by μ_n the measure $\mu(\{n\})$ of the set consisting of the single point n.) Moreover, $\mu \in \mathcal{E}$ if and only if $\sum(|\mu_n|+|\mu_{n+1}|)^2<+\infty$; or equivalently if and only if $\sum|\mu_n|^2<+\infty$. If $\mu\in\mathcal{E}$ and $\|\mu\|=0$, then $|\mu_n|$ is constant and hence equal to 0, i.e., $\mu=0$. Consequently, k is *strictly definite*, and hence *perfect*. If the potential $k*\mu$ of some measure $\mu\in\mathcal{E}$ vanishes outside some finite interval $a\leqslant n\leqslant b$, then μ is supported by the interval $a+1\leqslant n\leqslant b-1$. In fact, μ satisfies the difference equation

$$\mu_{n-1}+2\,\mu_n+\mu_{n+1}=0$$

in each of the regions $n<a$ and $n>b$; and hence $(-1)^n\mu_n$ is linear in the regions $n\leqslant a$ and $n\geqslant b$. Since $\sum|\mu_n|^2<+\infty$, we conclude that, actually, $\mu_n=0$ when $n\leqslant a$ or $n\geqslant b$. Denoting by ν the measure defined by $\nu_n=(-1)^n$, we obtain

$$\int (k*\mu)\,d\nu=\sum_n(\mu_{n-1}+2\,\mu_n+\mu_{n+1})\cdot(-1)^n=0,$$

the sum on the right being actually finite. This shows that the class of all potentials $k*\mu$ of compact support and with $\mu\in\mathcal{E}$ is *not rich*. Observe that $k*\nu=0$ and $j*\nu=0$, and hence, formally, $\|\nu\|^2=0$; but the energy of ν is not defined.

Example 10. As an example of a closed set which is *not capacitable* with respect to a perfect convolution kernel one may take the diagonal $D=\{(t,u):t\in T,\ u\in U,\ t=u\}$ in the product space $X=T\times U$, where T denotes the Abelian group R^3 with the discrete topology, and U denotes the same abstract group R^3, but with the usual Euclidean topology. Clearly, X is a locally compact Abelian group which is not of class K_σ. The kernel function k is defined as the tensor product $k=g\otimes h$ of the characteristic function g associated with the origin in T and the Newtonian kernel function h on U. Explicitly,

$$k(x)=k(t,u)=\begin{cases}|u|^{-1} & \text{for } t=0\\ 0 & \text{for } t\neq 0.\end{cases}$$

Since $|u|^{-1}=c\cdot|u|^{-2}*|u|^{-2}$ in U (c being a suitable constant, cf. § 8.1), and $g=g*g$ in T, we find that k has the form considered in Theorem 7.4, and k is therefore consistent. It is easily verified that k is strictly definite, and hence *perfect*. Denoting he interior Newtonian capacity of a subset $A\subset U$ by $\gamma_*(A)$, one may easily prove that the interior capacity with respect to k of an arbitrary set $E\subset X$ is determined by

$$\mathrm{cap}_* E = \sum_{t \in T} \gamma_* (E_t),$$

where $E_t = \{u \in U: (t, u) \in E\}$. For the diagonal D we find $\mathrm{cap}_* D = 0$ because D_t reduces to the single point $u = t$ whose Newtonian capacity is 0. On the other hand, $\mathrm{cap}^* D = +\infty$ because $\mathrm{cap}_* G = +\infty$ for every open set $G \supset D$. In fact, G_t contains for every t some neighbourhood of t, and hence $\gamma_* (G_t) > 0$.

References

[1]. N. Aronszajn & K. T. Smith, Functional spaces and functional completion. *Ann. Inst. Fourier Grenoble*, 6 (1955–56), 125–185.

[2]. N. Bourbaki, *Topologie générale*. Chap. I–II. Actualités Sci. Ind., 1142. Paris, 1951.

[3]. ——, *Topologie générale*. Chap. IX. Actualités Sci. Ind., 1045, Paris, 1958.

[4]. ——, *Intégration*. Chap. I–IV. Actualités Sci. Ind., 1175. Paris, 1952.

[5]. ——, *Intégration des mesures*. Actualités Sci. Ind., 1244. Paris, 1956.

[6]. ——, *Espaces vectoriels topologiques*. Chap. III–IV. Actualités Sci. Ind., 1229. Paris, 1955.

[7]. M. Brelot, La théorie moderne du potential. *Ann. Inst. Fourier Grenoble*, 4 (1952), 113–140.

[8]. H. Cartan, Sur la mesure de Haar. *C. R. Acad. Sci. Paris*, 211 (1940), 759–762.

[9]. ——, Sur les fondements de la théorie du potentiel. *Bull. Soc. Math. France*, 69 (1941), 71–96.

[10]. ——, Théorie du potentiel Newtonien. *Bull. Soc. Math. France*, 73 (1945), 74–106.

[11]. H. Cartan & J. Deny, Le principe du maximum en théorie du potentiel et la notion de fonction surharmonique. *Acta Sci. Math. Szeged*, 12 (1950), 81–100.

[12]. G. Choquet, Les noyeaux réguliers en théorie du potentiel. *C. R. Acad. Sci. Paris*, 243 (1956), 635–638.

[13]. ——, Sur les fondements de la théorie fine du potentiel. *C. R. Acad. Sci. Paris*, 244 (1957). 1606–1609.

[14]. ——, Theory of capacities. *Ann. Inst. Fourier Grenoble*, 5 (1953–54), 131–295.

[15]. J. Deny, Les potentiels d'énergie finie. *Acta Math.*, 82 (1950), 107–183.

[16]. ——, Sur la définition de l'énergie en théorie du potentiel. *Ann. Inst. Fourier Grenoble*, 2 (1950), 83–99.

[17]. G. C. Evans, On potentials of positive mass. I. *Trans. Amer. Math. Soc.*, 37 (1935), 226–253.

[18]. O. Frostman, Potentiel d'équilibre et capacité des ensembles. *Comm. Sém. Math. Univ. Lund*, 3 (1935).

[19]. M. Kishi, Capacities of Borelian sets and the continuity of potentials. *Nagoya Math. J.*, 12 (1957), 195–219.

[20]. K. Kunugui, Étude sur la théorie du potentiel généralisé. *Osaka Math. J.*, 2 (1950), 63–103.

[21]. M. A. J. Maria, The potential of a positive mass and the weight function of Wiener. *Proc. Nat. Acad. Sci. U.S.A.*, 20 (1934), 485–489.

[22]. N. Ninomiya, Equilibrium potentials and energy integrals. *Proc. Japan Acad.*, 26 (1950), 1–16.

[23]. ——, Etude sur la théorie du potentiel pris par rapport à un noyau symétrique. *J. Inst. Polytech., Osaka City Univ. Ser. A. Math.*, 8 (1957), 147–179.

[24]. M. OHTSUKA, Sur un espace complet de mesures positives dans la théorie du potentiel. *Proc. Japan Acad.*, 32 (1956), 311–313.

[25]. ——, Les relations entre certains principes en théorie du potentiel. *Proc. Japan Acad.*, 33 (1957), 37–40.

[26]. M. RIESZ, Intégrales de Riemann-Liouville et potentiels. *Acta Sci. Math. Szeged*, 9 (1938), 1–42.

[27]. L. SCHWARTZ, *Théorie des distributions.* I–II. Paris, 1950–51.

[28]. T. UGAHERI, On the general capacities and potentials. *Bull Tokyo Inst. Tech. Ser. B.*, 4 (1953), 149–179.

[29]. Ch. de la VALLÉE-POUSSIN, *Les nouvelles méthodes de la théorie du potentiel* ... . Actualités Sci. Ind., 578. Paris, 1937.

[30]. ——, *Le potentiel logarithmique.* Paris-Louvain, 1949.

[31]. F. VASILESCO, Sur la continuité du potentiel à travers des masses. *C. R. Acad. Sci. Paris,* 200 (1935), 1173–1174.

[32]. D. V. WIDDER, *The Laplace Transform.* Princeton, 1946.

[33]. N. WIENER, Certain notions in potential theory. *J. Math. Physics,* 3 (1924), 24–51.

[34]. K. YOSIDA, Sur le principe du maximum dans la théorie du potentiel. *Proc. Japan Acad.*, 17 (1941), 476–478.

Received June 5, 1959.