Journal of Statistical Physics, Vol. 61, Nos. 1/2, 1990

On Ruch's Principle of Decreasing Mixing Distance in Classical Statistical Physics

Paul Busch[1] and Ralf Quadt[1]

Received December 29, 1989; final May 11, 1990

Ruch's Principle of Decreasing Mixing Distance is reviewed as a statistical physical principle and its basic suport and geometric interpretation, the Ruch–Schranner–Seligman theorem, is generalized to be applicable to a large representative class of classical statistical systems.

KEY WORDS: Irreversibility; stochastic operator; mixing character; mixing distance.

1. INTRODUCTION

With the concept of mixing distance Ernst Ruch has opened a new approach to the phenomena of irreversibility and organization in the physical sciences. The mixing distance is a generalization of the mixing character, which was originally introduced as a (partial) ordering on the set of partitions of a finite set (the *diagram lattice*),[1,2] but was soon generalized to apply to spaces of distributions. Natural fields of application are the chirality and isomery phenomena in chemistry,[1] the theory of representations of the symmetric group,[2] graph theory,[3] and statistical thermodynamics.[4–7] Considering the enormous scope of the notions of mixing character and mixing distance, one is readily convinced of their fundamental nature, which also becomes apparent in their deep geometrical meaning as elucidated very recently by Ruch.[8] We therefore feel encouraged to accept the physical importance of Ruch's concept and to explore its implications.

The irreversible evolution of complex systems toward some equilibrium state or toward some configuration of organization is generally described in

---

This paper is dedicated to Ernst Ruch on the occasion of his 70th birthday.

[1] Institut für Theoretische Physik, Universität zu Köln, D-5000 Köln 41, Federal Republic of Germany.

terms of an increase of entropy, or relative entropy (whenever an entropy can be defined). Ruch proposed to sharpen the criterion of (relative) entropy increase into the criteria of increasing mixing character, resp. decreasing mixing distance, thus strengthening the *second law* of thermodynamics for closed, resp. open, systems. The motivation for this *Principle of Decreasing Mixing Distance* is twofold. First, it indeed implies increase of (relative) entropy. Second, if one accepts the (linear) generalized master equation[9] as the general statistical description of the dynamics of macroscopic systems, then decreasing mixing distance (and not only increasing entropy) is a consequence. An even more stringent argument in favor of this principle is the reversed statement formulated in the Ruch–Schranner–Seligman (RSS) theorem.[10] It is the purpose of the present paper to generalize this theorem so that it becomes applicable to a reasonably large class of statistical systems.

We shall first review the Principle of Decreasing Mixing Distance and its foundations, the RSS theorem (Section 2). Next the generalized RSS theorem is formulated and proved (Section 3). Finally, as an example, the discrete case of that theorem is revisited and slightly extended (Section 4) and fields of potential applications are briefly outlined (Section 5).

2. THE PRINCIPLE OF DECREASING MIXING DISTANCE

Rather than following the chronological development of the concept of mixing distance, we shall sketch a systematic, deductive presentation. We report very briefly some geometrical features and strongly recommend Ruch's fundamental work[8] as the original and thorough treatment. Following Ruch,[8] the mixing distance is a special instance of the direction distance, the latter being a formalization of the geometric concept of angle in real vector spaces with an arbitrary (not necessarily Hilbert type) norm. In short, an *oriented angle* is defined as an ordered pair $([x], [y])$ of directions ($[x] = \{\alpha x \mid \alpha > 0\}$ for $x \neq 0$). Generalizing Felix Klein's concept of *congruence*, two geometrical figures shall be called *norm equivalent* if there exists a linear transformation which maps one figure bijectively and isometrically onto the other one. Thus, pairs of oriented angles $([x], [y])$, $([x'], [y'])$ are norm equivalent if they obey the following relations:

$$\|\alpha x_0 - \beta y_0\| = \|\alpha x'_0 - \beta y'_0\| \qquad \forall \alpha, \beta \in \mathbb{R}^+$$

Here z_0 means $z/\|z\|$ for $z \neq 0$. Now one obtains a (partial) ordering of the equivalence classes of oriented angles, denoted $d[x/y] \succ d[x'/y']$, via the system of inequalities

$$\|\alpha x_0 - \beta y_0\| \geqslant \|\alpha x'_0 - \beta y'_0\| \qquad \forall \alpha, \beta \in \mathbb{R}^+$$

The entity $d[x/y]$—the *direction distance* (of x from y)—may be identified with the function

$$d[x/y]\colon\quad \mathbb{R}^+ + \mathbb{R}^+ \to \mathbb{R}^+$$
$$(\alpha, \beta) \mapsto d[x/y](\alpha, \beta) = \|\alpha x_0 - \beta y_0\|$$

As Ruch[8] has shown, the direction distance has properties quite similar to those of a metric, which justifies the use of the term "distance." In fact, one readily verifies the following relations, which are in complete analogy to the distance axioms:

(d_1) $d[x/y] \succ d[z/z]$; $d[x/y] = d[z/z]$ iff $[x] = [y]$.

(d_2) $d[x/y] \succ d[x'/y']$ iff $d[y/x] \succ d[y'/x']$ ("symmetry").

(d_3) $d[x/y] \prec d[x/z] + d[z/z] + d[z/y]$ ("triangle inequality").

Note that $d[z/z]$ plays the role of the null distance. It is crucial that the direction distance is not a mere number, but a function: this is a reflection of the fact that the above ordering of angles is no *total* ordering. It becomes a total ordering exactly if the underlying norm induces an inner product (via the polarization identity); this happens if and only if the familiar symmetry property holds: $d[x/y] = d[y/x]$. In this way the familiar notion of angle for inner product (pre-Hilbert) spaces is recovered exactly if there is a group of linear automorphisms acting transitively on the norm unit sphere. In other words, the direction distance is the canonically generalized notion of a metric for (oriented) angles in the context of affine geometries.

The direction distance is known under the name *mixing distance* in the case of a particular family of vector spaces relevant to statistical physics, namely vector spaces $M = M(\Omega, \Sigma)$ of bounded σ-additive signed measures on a measurable space (Ω, Σ). The set of all (positive) measures forms a proper convex cone M^+ in M generating M (such that $M = M^+ - M^+$). Furthermore, the normalized (probability) measures form a convex subset $S = M_1^+$ of M^+ and as such part of a hyperplane of M; S uniquely determines a positive linear functional e (the "trace," or "charge" functional):

$$e\colon\quad M \to \mathbb{R}$$
$$x \mapsto e(x) := x(\Omega)$$

such that

$$S = \{x \in M^+ \mid e(x) = x(\Omega) = 1\}$$

[We note that positivity of e means $e(M^+) \subseteq \mathbb{R}^+$.] Now M becomes a normed space in a canonical way by considering the convex hull of $S \cup -S$,

$$M_1 = \operatorname{conv}(S \cup -S) = \{\alpha x - \beta y \mid x, y \in S, \alpha, \beta \geqslant 0, \alpha + \beta = 1\}$$

as the norm unit ball. The resulting norm is the Minkowski functional of M_1 and satisfies

$$x \mapsto \|x\|_1 = \sup\{x(A) \mid A \in \Sigma\} - \inf\{x(A) \mid A \in \Sigma\} = |x|(\Omega)$$

(where $|x|$ denotes the total variation of the signed measure x). It follows that on M^+ the norm $\|\cdot\|_1$ coincides with the trace e. We note that there exists a family of "measure-type" norms, including $\|\cdot\|_1$, all yielding the same equivalence and ordering relations for the oriented angles $d[p/q]$ between *positive* elements $p, q \in M^+$ (measures!); this again emphasizes the relevance of the mixing distance in the measure-theoretic and therefore also in the statistical context.[11] In the following we restrict our attention to the norm $\|\cdot\|_1$ given by the total variation (occasionally called 1-norm).

The physical importance of the spaces $M = M(\Omega, \Sigma)$ is due to the possibility of interpreting measures from M^+ as distributions and normalized measures from $S = M_1^+$ as probability measures. In the statistical context the set S of probability measures is referred to as the set of (statistical) *states*, defined on the *phase space* Ω. A statistical state represents an ensemble of identically prepared macrosystems. An alternative interpretation refers to the elements of S as describing the (average) *distribution* of macroscopic subsystems of a larger system over a reduced phase space (example: the Boltzmann one-particle distributions). In the sequel we shall refer to the first interpretation. The restriction of the direction distance $d[p/q]$ to pairs p, q from M_1^+ is called the mixing distance. With respect to the geometrical interpretation, it is important to observe that there exists a semigroup of linear mappings Φ on M leaving S invariant $[\Phi(S) \subseteq S]$ and acting transitively on S. These linear mappings are positive $[\Phi(M^+) \subset M^+]$, trace-preserving $(e \circ \Phi = e)$, and therefore contractive $(\|\Phi x\|_1 \leqslant \|x\|_1,\ x \in M)$; in the statistical context they are known as *stochastic operators* (linear state transformations). Henceforth the set of stochastic operators on M shall be denoted $\mathrm{ST}(M)$. Being contractions, stochastic operators lead to decreasing mixing distance, that is, if $p' = \Phi p$ and $q' = \Phi q$ $(p, q \in S)$, then $d[p/q] \succ d[p'/q']$. The RSS theorem states basically the converse:

Theorem 2.1 (Ruch, Schranner, Seligman[10]). Let $M = M(\Omega, \Sigma)$ be one of the following normed spaces of finite signed measures:

1. $M = M(\Omega, \Sigma) \cong \mathbb{R}^n$ for $\Omega = \{\omega_1, \ldots, \omega_n\}$, $\Sigma = \mathfrak{P}(\Omega)$.

2. $M=M(\Omega, \Sigma) \cong L^1([0,1], dx)$ for $\Omega=[0,1]$, $\Sigma=\mathfrak{B}(\Omega)$, the Borel algebra on the space Ω, $dx=$ Lebesgue measure.

Then the following statements for $p, q, p', q' \in S=M_1^+$ are equivalent:

(i) $d[p/q] \succ d[p'/q']$.

(ii) There exists a stochastic operator $\Phi \in \mathrm{ST}(M)$ such that $p'=\Phi p$ and $q'=\Phi q$.

In the proof, case 1 is considered as entailed by case 2. We shall return to this point in Section 4.

It can be reasonably expected that the validity of the RSS theorem extends to more general spaces $M(\Omega, \Sigma)$. Indeed, in Section 3 we shall provide a generalization which makes the theorem applicable to a wide class of statistical physical systems.

It is worthwhile recalling the special case of Theorem 2.1 referring to the mixing character. In the above two cases the space $M=M(\Omega, \Sigma)$ is such that it contains an element representing the *uniform* distribution $u \in M_1^+$; then the *mixing character* may be defined as

$$m[p]=-d[p/u]$$

The entity $m[p]$ represents a qualification of the similarity of a distribution p to the uniform distribution in the sense of a (partial) ordering

$$m[p'] \succ m[p] \quad \text{iff} \quad \|p-lu\|_1 \geqslant \|p'-lu\|_1 \qquad \forall l \geqslant 0$$

A stochastic operator Φ is called *bistochastic* if the uniform distribution u is an eigenvector: $\Phi u=u$. Then the RSS theorem entails the Hardy–Littlewood–Polya (HLP) theorem:

Theorem 2.2 (Hardy, Littlewood, Polya/Ryff). Let M be as in Theorem 2.1. Then the following statements for $p, p' \in S$ are equivalent:

(i) $m[p'] \succ m[p]$.

(ii) There exists a bistochastic operator Φ on M such that $p'=\Phi p$.

These theorems show the deep connection between the notion of angle in spaces of signed measures and the concept of stochastic operators. As mentioned in the introduction, the latter concept is fundamental to the general description of *irreversible* (stochastic) dynamics. In fact, in the usual description of classical statistical systems, observables are represented as (essentially) bounded functions on phase space Ω, that is, as elements of some space $L^\infty(\Omega, \mu)$ of μ-essentially bounded functions on Ω, where (Ω, Σ, μ) is the underlying event space, μ some measure on (Ω, Σ). [A more systematic approach which allows for a unified description of classical and quantum statistical systems identifies observables as (normalized) measures on (Ω, Σ) with range in the positive elements of $M(\Omega, \Sigma)'$.] Accordingly, taking into account the finite resolution of any measurement, statistical states are represented as absolutely continuous (with respect to μ) probability measures on (Ω, μ), that is, as positive normalized elements from $M_\mu(\Omega) \cong L^1(\Omega, \mu)$. In Section 3 a class of physically relevant measure spaces (Ω, Σ, μ) will be specified to which the RSS theorem can be extended. Thus, the RSS theorem establishes a geometric interpretation of the dynamics (via stochastic operators) of macroscopic systems, which is summarized in Ruch's Principle.

Principle of Decreasing Mixing Distance. The evolution of a physical system always leads to a decrease of mixing distance with increasing time. The evolution of a closed system always leads to increasing mixing character.

This Principle assumes the status of the second law insofar as it gives rise to a general characterization of irreversibility. The second law of thermodynamics follows if one takes into account that the inequalities $d[p/q] \succ d[p'/q']$ (resp. $m[p'] \succ m[p]$) are equivalent to systems of inequalities $f(p, q) \geqslant f(p', q')$ [resp. $g(p') \geqslant g(p)$] for certain classes $\mathfrak{MD}$ (resp. $\mathfrak{MC}$) of convex functions in (p, q) (resp. p) *containing relative entropy* (resp. entropy).[4],[2] This equivalence furnishes the power of Ruch's Principle: it is obvious that these systems of inequalities entail much more information about (or impose stronger restrictions on) the possible dynamics on statistical state spaces than is provided by entropy alone.

3. EXTENSION OF THE RUCH–SCHRANNER–SELIGMAN THEOREM

In this section we shall extend Theorem 2.1 to a wide class of measure spaces (Ω, Σ, μ). It will be assumed that (Ω, Σ, μ) is a separable measure space with a σ-finite measure μ. The proof of our main result will be based on some facts from the theory of C^*-algebras, which, for the sake of convenience, will be explained in some detail.

We shall also sketch an alternative proof of the theorem by means of purely measure-theoretic means. To spell out all details would require a roughly equal amount of space. The reason for our preferring the algebraic approach is that this procedure first is closer to the language of statistical

[2] We note that results of the general form of Theorem 2.1 (for n-tuples of states on an arbitrary commutative C^*-algebra) were derived by Alberti and Uhlmann,[12] who, instead of our condition (i), require a set of inequalities for a family $\mathfrak{MD}$ of "h-convex" functionals. For the case of *pairs* of states our result is stronger, since it refers to a *minimal* set of h-convex functionals.

physics and second paves the way for extensions of the result to quantum statistical systems (to be treated in terms of noncommutative operator algebras).

The extended RSS theorem states the following:

Theorem 3.1. Let $f, g, f', g' \in L^1(\Omega, \Sigma, \mu)_1^+$, where (Ω, Σ, μ) is a separable, σ-finite measure space. Then the following are equivalent:

(i) $d[f/g] \succ d[f'/g']$, $d[f/g]\colon \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$, $(\alpha, \beta) \mapsto \|\alpha f - \beta g\|_1$.

(ii) There exists a stochastic linear mapping $\Phi \in \mathrm{ST}(L^1(\Omega, \Sigma, \mu))$ such that $f' = \Phi f$ and $g' = \Phi g$.

Proof. The implication (ii) $\Rightarrow$ (i) is obvious. To prove the converse, we restrict ourselves for the present to the case of atom-free, separable, σ-finite measure spaces. A measure space is called *atom-free* if the quotient algebra $\mathfrak{B} \equiv \Sigma/\Delta_\mu$ (the set of equivalence classes $[E]_{\Delta_\mu}$ in Σ with respect to the σ-ideal $\Delta_\mu = \{A \in \Sigma \mid \mu(A) = 0\}$, $[E]_{\Delta_\mu} = \{A \in \Sigma \mid A\, \Delta E := (A - E) \cup (E - A) \in \Delta_\mu\}$) contains no atoms. We make use of the fact that the dual spaces $L^\infty_{\mathbb{C}}(\Omega, \Sigma, \mu)$ and $L^\infty_{\mathbb{C}}([0, 1], dx)$ of the complex Banach spaces $L^1_{\mathbb{C}}(\Omega, \Sigma, \mu)$ and $L^1_{\mathbb{C}}([0, 1], dx)$ are W^*-algebras which are isomorphic to each other in the case of atom-free measure spaces. (Henceforth we shall omit the subscript $\mathbb{C}$, as it is evident from the context whether the complex or the real spaces are referred to.) In subsequent lemmas we shall provide an explicit construction of such an isomorphism which, together with its inverse, is normal. This enables us to deduce the implication (i) $\Rightarrow$ (ii) by application of the corresponding implication of Theorem 2.1.

Let $i\colon L^\infty(\Omega, \Sigma, \mu) \to L^\infty([0, 1], dx)$ be a normal isomorphism and i^{-1} its normal inverse. Then the transposed mappings i^t and $(i^{-1})^t$ are positive, linear isometries satisfying $i^t(L^1([0, 1], dx)) \subset L^1(\Omega, \Sigma, \mu)$ and $(i^{-1})^t\,(L^1(\Omega, \Sigma, \mu)) \subset L^1([0, 1], dx)$. [Here, as well as in the following, we implicitly consider the space $L^1(\Omega, \Sigma, \mu)$ as embedded into its second dual in the canonical way.] Now assume (i): $\|\alpha f - \beta g\|_1 \geqslant \|\alpha f' - \beta g'\|_1$ $\forall \alpha, \beta \geqslant 0$. Then for $\tilde{f} = (i^{-1})^t f$, $\tilde{g} = (i^{-1})^t g$, $\tilde{f}' = (i^{-1})^t f'$, and $\tilde{g}' = (i^{-1})^t g'$ (all in $L^1([0, 1], dx)$), one has $\alpha\tilde{f} - \beta\tilde{g} = (i^{-1})^t(\alpha f - \beta g)$ and $\alpha\tilde{f}' - \beta\tilde{g}' = (i^{-1})^t(\alpha f' - \beta g')$ and therefore [since $(i^{-1})^t$ is an isometry]

$$\|\alpha\tilde{f} - \beta\tilde{g}\|_1 \geqslant \|\alpha\tilde{f}' - \beta\tilde{g}'\|_1 \qquad \forall \alpha, \beta \geqslant 0$$

According to Theorem 2.1, there exists a $\tilde{\Phi} \in \mathrm{ST}(L^1([0, 1], dx))$ such that $\tilde{f}' = \tilde{\Phi}\tilde{f}$ and $\tilde{g}' = \tilde{\Phi}\tilde{g}$. But this yields $f' = i^t\tilde{f}' = [i^t\tilde{\Phi}(i^{-1})^t](f)$ and $g' = [i^t\tilde{\Phi}(i^{-1})^t](g)$, which means $f' = \Phi f$ and $g' = \Phi g$ with $\Phi \in \mathrm{ST}(L^1(\Omega, \Sigma, \mu))$.

The extension of this result to the case of measure spaces containing atoms will be based on Lemma 3.6, which states that one can embed a separable, σ-finite measure space (Ω, Σ, μ) into an atom-free, separable, σ-finite measure space $(\tilde{\Omega}, \mathfrak{F}, \tilde{\mu})$. Thus, it is possible to apply the isometries i^t, $(i^{-1})^t$ which guarantee the existence of a stochastic operator $\Phi \in \mathrm{ST}(L^1(\tilde{\Omega}, \mathfrak{F}, \tilde{\mu}))$. By the second part of Lemma 3.6, Φ induces a stochastic operator $\Phi' \in \mathrm{ST}(L^1(\Omega, \Sigma, \mu))$. In this way the validity of the implication (i) $\Rightarrow$ (ii) is transferred from $L^1(\tilde{\Omega}, \mathfrak{F}, \tilde{\mu})$ to $L^1(\Omega, \Sigma, \mu)$. QED

Now we shall proceed to construct a normal isomorphism i between the W^*-algebras $L^\infty(\Omega, \Sigma, \mu)$ and $L^\infty([0, 1], dx)$. For this purpose we need the following:

Lemma 3.2. Let (Ω, Σ, μ) be a σ-finite measure space. Then there exists a faithful representation $(\pi, L^2(\Omega, \Sigma, \mu))$ of the W^*-algebra $L^\infty(\Omega, \Sigma, \mu)$, that is, a $*$-isomorphism π from $L^\infty(\Omega, \Sigma, \mu)$ onto an Abelian von Neumann algebra $\mathscr{A}$ in $\mathscr{L}(L^2(\Omega, \Sigma, \mu))$. Furthermore, π can be chosen bicontinuous with respect to the $\sigma(L^\infty(\Omega, \Sigma, \mu), L^1(\Omega, \Sigma, \mu))$- and the σ-weak [i.e., the $\sigma(\mathscr{A}, \mathscr{L}(L^2(\Omega, \Sigma, \mu))_*)$-] topologies.

Proof. To every $\varphi \in L^\infty(\Omega, \Sigma, \mu)$ an operator $M_\varphi: L^2(\Omega, \Sigma, \mu) \to L^2(\Omega, \Sigma, \mu)$ can be assigned which is uniquely determined by the (μ-almost everywhere) pointwise equality $M_\varphi f = \varphi f$ for $f \in L^2(\Omega, \Sigma, \mu)$. Thus, one can define a mapping

$$\pi: \quad L^\infty(\Omega, \Sigma, \mu) \to \mathscr{A} = \{M_\varphi \mid \varphi \in L^\infty(\Omega, \Sigma, \mu)\}$$
$$\varphi \mapsto \pi(\varphi) = M_\varphi$$

Obviously, $\mathscr{A} \subset \mathscr{L}(L^2(\Omega, \Sigma, \mu))$, π is a bijection satisfying:

(i) $\pi(\alpha\varphi + \beta\psi) = \alpha\pi(\varphi) + \beta\pi(\psi)$, $\forall \varphi, \psi \in L^\infty(\Omega, \Sigma, \mu)$, $\forall \alpha, \beta \in \mathbb{C}$.

(ii) $\pi(\varphi\psi) = \pi(\varphi)\,\pi(\psi)$, $\forall \varphi, \psi \in L^\infty(\Omega, \Sigma, \mu)$.

(iii) $\pi(\bar{\varphi}) = \pi(\varphi)^*$, $\forall \varphi \in L^\infty(\Omega, \Sigma, \mu)$.

Properties (i) and (ii) are obvious, and (iii) follows from

$$(M_\varphi f, g) = \int (\varphi f)\,\bar{g}\,d\mu = \int f\,\overline{(\bar{\varphi} g)}\,d\mu = (f, M_{\bar{\varphi}}\,g)$$

Thus $(\pi, L^2(\Omega, \Sigma, \mu))$ is a $*$-representation of the W^*-algebra $L^\infty(\Omega, \Sigma, \mu)$. Furthermore, π is strictly positive, which implies $\ker \pi = \{0\}$ and finally $\|\pi(\varphi)\| = \|\varphi\|_\infty$ for all $\varphi \in L^\infty(\Omega, \Sigma, \mu)$. Consequently, π is a $*$-isomorphism.

Next we shall prove that π is a normal representation $(\pi'(\mathscr{L}(L^2(\Omega, \Sigma, \mu))_*) \subset L^1(\Omega, \Sigma, \mu))$, i.e., π is σ-weakly continuous. Then Proposition III.3.12 of ref. 13 ensures that $\mathscr{A} = \pi(L^\infty(\Omega, \Sigma, \mu))$ is a von Neumann algebra. Let $\{\varphi_\alpha\} \subset L^\infty(\Omega, \Sigma, \mu)$ be a $\sigma(L^\infty(\Omega, \Sigma, \mu), L^1(\Omega, \Sigma, \mu))$-converging net, that is, there exists $\varphi \in L^\infty(\Omega, \Sigma, \mu)$ such that $\langle f, \varphi_\alpha - \varphi \rangle \to 0$ for all $f \in L^1(\Omega, \Sigma, \mu)$. Then one has to prove that $\langle \rho, \pi(\varphi_\alpha) - \pi(\varphi) \rangle \to 0$ for all $\rho \in \mathscr{L}(L^2(\Omega, \Sigma, \mu))_*$. For each $\rho \in \mathscr{L}(L^2(\Omega, \Sigma, \mu))_*$, there exist sequences $\{\xi_n\}$ and $\{\eta_n\}$ in $L^2(\Omega, \Sigma, \mu)$ with $\sum_{n=1}^\infty \|\xi_n\|_2^2 < +\infty$ and $\sum_{n=1}^\infty \|\eta_n\|_2^2 < +\infty$ such that ρ has the form $\rho = \sum_{n=1}^\infty \omega_{\xi_n, \eta_n}$. [Here $\omega_{\xi,\eta}$ is the linear functional given by $\omega_{\xi,\eta}(a) = (\xi, a\eta)$ for a in $\mathscr{L}(L^2(\Omega, \Sigma, \mu))$.] Therefore, applying the Lebesgue dominated convergence theorem,

$$\begin{aligned}
&\langle \rho, \pi(\varphi_\alpha) - \pi(\varphi) \rangle \\
&\quad = \sum_{n=1}^\infty (\xi_n, (M_{\varphi_\alpha} - M_\varphi)\eta_n) \\
&\quad = \int_\Omega d\mu \left(\sum_{n=1}^\infty \bar{\xi}_n(x)\, \eta_n(x) \right) [\varphi_\alpha(x) - \varphi(x)] \\
&\quad = \int_\Omega d\mu\, g(x)[\varphi_\alpha(x) - \varphi(x)] \\
&\quad = \langle g, \varphi_\alpha - \varphi \rangle \to 0
\end{aligned}$$

which converges to zero, since the function g defined by $g(x) = \sum_{n=1}^\infty \bar{\xi}_n(x)\, \eta_n(x)$ is in $L^1(\Omega, \Sigma, \mu)$:

$$\begin{aligned}
\int_\Omega |g|\, d\mu &\leqslant \sum_{n=1}^\infty (|\xi_n|, |\eta_n|) \leqslant \sum_{n=1}^\infty \|\xi_n\|_2\, \|\eta_n\|_2 \\
&\leqslant \left(\sum_{n=1}^\infty \|\xi_n\|_2^2 \right)^{1/2} \left(\sum_{n=1}^\infty \|\eta_n\|_2^2 \right)^{1/2} < +\infty
\end{aligned}$$

Here we made use of the Cauchy–Schwartz inequality for $L^2(\Omega, \Sigma, \mu)$ and l^2.

Finally, we have to prove that also π^{-1} is σ-weakly continuous, i.e., $(\pi^{-1})^t (L^1(\Omega, \Sigma, \mu)) \subset \mathscr{L}(L^2(\Omega, \Sigma, \mu))_*$. Let $\{M_{\varphi_\alpha}\} \subset \mathscr{A}$ be a net which converges σ-weakly to $M_\varphi \in \mathscr{A}$: $\langle \rho, M_{\varphi_\alpha} - M_\varphi \rangle \to 0$ for all $\rho \in \mathscr{L}(L^2(\Omega, \Sigma, \mu))_*$. We have to show that $\langle f, \pi^{-1}(M_{\varphi_\alpha}) - \pi^{-1}(M_\varphi) \rangle \to 0$ for any $f \in L^1(\Omega, \Sigma, \mu)$. Every function $f \in L^1(\Omega, \Sigma, \mu)$ can be represented as $f(x) = \bar{\xi}(x)\, \eta(x)$ with $\xi(x)$, $\eta(x) \in L^2(\Omega, \Sigma, \mu)$. Thus, for each $f \in L^1(\Omega, \Sigma, \mu)$ there exists a functional $\omega_{\xi,\eta} \in \mathscr{L}(L^2(\Omega, \Sigma, \mu))_*$ such that $\langle f, \psi \rangle = \langle \omega_{\xi,\eta}, \pi(\psi) \rangle$ for all $\psi \in L^\infty(\Omega, \Sigma, \mu)$. Therefore, we have $\langle f, \varphi_\alpha - \varphi \rangle = \langle \omega_{\xi,\eta}, M_{\varphi_\alpha} - M_\varphi \rangle \to 0$. QED

Next we shall make use of the fact that the measure space (Ω, Σ, μ) is supposed to be atom-free.

Lemma 3.3. Let (Ω, Σ, μ) be an atom-free, σ-finite measure space. Then the von Neumann algebra $\mathscr{A} = \pi(L^\infty(\Omega, \Sigma, \mu)) \subset \mathscr{L}(L^2(\Omega, \Sigma, \mu))$ given by Lemma 3.2 contains no minimal projections (atoms).

Proof. We shall prove that the W^*-algebra $L^\infty(\Omega, \Sigma, \mu)$ contains no minimal projection. This implies that $\mathscr{A}$ contains no minimal projection for the following reason. The $*$-isomorphism π of Lemma 3.2 induces a lattice isomorphism between $\mathscr{P}(\mathscr{A})$ and $\langle \mathscr{X}, \wedge, \neg, 1\rangle$, the projection lattice of $\mathscr{A}$ and the complete lattice of projections in $L^\infty(\Omega, \Sigma, \mu)$. A lattice isomorphism is order preserving in both directions, and therefore $L^\infty(\Omega, \Sigma, \mu)$ contains no atoms if, and only if, $\mathscr{A}$ contains no atoms.

Let $\Delta_\mu = \{S \in \Sigma \mid \mu(S) = 0\}$. The classes $[A]_{\Delta_\mu} := \{B \in \Sigma \mid A \Delta B \equiv A - B \cup B - A \in \Delta_\mu\}$ form a Boolean algebra (the quotient algebra $\mathfrak{B} = \{[A]_{\Delta_\mu} \mid A \in \Sigma\} = \Sigma/\Delta_\mu$ under the following Boolean lattice operations: $\neg [A]_{\Delta_\mu} = [\Omega - A]_{\Delta_\mu}$, $[A]_{\Delta_\mu} \cap [B]_{\Delta_\mu} = [A \cap B]_{\Delta_\mu}$, and $[A]_{\Delta_\mu} \cup [B]_{\Delta_\mu} = [A \cup B]_{\Delta_\mu}$. By assumption, the Boolean lattice $\langle \mathfrak{B}, \cap, \neg, 1\rangle$ contains no atoms.

Now let

$$\chi_{[E]_{\Delta_\mu}} := \{\chi_A \mid A \in [E]_{\Delta_\mu}\} \equiv \bar{\chi}_E$$

(χ_A denotes the characteristic function of A) and

$$\mathscr{X} := \{\chi_{[E]_{\Delta_\mu}} \mid [E]_{\Delta_\mu} \in \mathfrak{B}\}$$

The set $\mathscr{X}$ is a Boolean algebra under the following lattice operations: $\bar{\chi}_A \wedge \bar{\chi}_B \equiv \bar{\chi}_A \bar{\chi}_B = \bar{\chi}_{A \cap B}$, $\bar{\chi}_A \vee \bar{\chi}_B \equiv \bar{\chi}_A + \bar{\chi}_B - \bar{\chi}_A \bar{\chi}_B = \bar{\chi}_{A \cup B}$, and $\neg\, \bar{\chi}_A \equiv 1 - \bar{\chi}_A = \bar{\chi}_{\Omega - A}$. Now the mapping $i: \mathfrak{B} \to \mathscr{X}$, $[E]_{\Delta_\mu} \mapsto \bar{\chi}_E$ is a lattice isomorphism, so that together with the lattice $\langle \mathfrak{B}, \cap, \neg, 1\rangle$, also the lattice $\langle \mathscr{X}, \wedge, \neg, 1\rangle$ contains no atoms (by the same arguments as above). But by construction $\mathscr{X}$ can be identified with the projection lattice of $L^\infty(\Omega, \Sigma, \mu)$, so that this algebra—and therefore $\mathscr{A}$—contains no minimal projections. QED

The following theorem presupposes the separability of the measure space (Ω, Σ, μ). [A measure space (Ω, Σ, μ) is called separable if the associated metric space $(\Sigma(\mu), \rho)$ is separable. Here $\Sigma(\mu)$ denotes the set of all elements of finite measure in Σ and the metric ρ on $\Sigma(\mu)$ is given by $\rho(E, F) = \mu(E \Delta F)$.] We recall that the Banach space $L^p(\Omega, \Sigma, \mu)$ $(p < \infty)$ is separable if, and only if, $(\Sigma(\mu), \rho)$ is separable (ref. 14, p. 177). This guarantees the separability of the Hilbert space $L^2(\Omega, \Sigma, \mu)$ provided the measure space (Ω, Σ, μ) is separable. This enables us to apply Theorem III.1.22 of ref. 13:

Theorem 3.4. Let $\mathscr{A}$ be an Abelian von Neumann algebra on a separable Hilbert space. If $\mathscr{A}$ contains no nonzero minimal projection, then $\mathscr{A}$ is isomorphic to the algebra $L^\infty(0, 1) = L^\infty([0, 1], dx)$ of all essentially bounded functions on the unit interval $(0, 1)$ with respect to the Lebesgue measure.

This theorem implies the following important result:

Corollary 3.5. Let (Ω, Σ, μ) be a σ-finite, atom-free, separable measure space. Then there exists an isomorphism between the W^*-algebras $L^\infty(\Omega, \Sigma, \mu)$ and $L^\infty(0, 1)$.

It remains to prove that the isomorphism $i: L^\infty(\Omega, \Sigma, \mu) \to L^\infty(0, 1)$ from Corollary 3.5 and its inverse i^{-1} are normal, i.e., $i'(L^1(0, 1)) \subset L^1(\Omega, \Sigma, \mu)$ and $(i^{-1})' (L^1(\Omega, \Sigma, \mu)) \subset L^1(0, 1)$. These statements can be derived from the following commutative diagram:

$$\begin{array}{ccc}
L^\infty(\Omega, \Sigma, \mu) & \underset{i^{-1}}{\overset{i}{\rightleftarrows}} & L^\infty(0, 1) \\
\pi \downarrow \uparrow \pi^{-1} & & \tilde{\pi} \downarrow \uparrow \tilde{\pi}^{-1} \\
\mathscr{A} & \underset{\tau^{-1}}{\overset{\tau}{\rightleftarrows}} & \tilde{\mathscr{A}} \\
\subset \mathscr{L}(L^2(\Omega, \Sigma, \mu)) & & \subset \mathscr{L}(L^2(0, 1))
\end{array}$$

$\pi, \tilde{\pi}$ are isomorphisms given as in Lemma 3.2 and i denotes the isomorphism from Corollary 3.5. Thus $\tau = \tilde{\pi} \circ i \circ \pi^{-1}: \mathscr{A} \to \tilde{\mathscr{A}}$ and $\tau^{-1} = \pi \circ i^{-1} \circ \tilde{\pi}^{-1}: \tilde{\mathscr{A}} \to \mathscr{A}$ are isomorphisms between the von Neumann algebras $\mathscr{A}$ and $\tilde{\mathscr{A}}$. Then Corollary III.3.10 of ref. 13 ensures the σ-weak bicontinuity of τ, τ^{-1}. As shown in Lemma 3.2, π and $\tilde{\pi}$ are bicontinuous and these facts imply that both i and i^{-1} are normal. This completes the construction of a normal isomorphism.

Now we shall sketch the purely measure-theoretic proof of Theorem 3.1 mentioned in the beginning of this section. This proof rests on the fact that an atom-free, separable, σ-finite measure space (Ω, Σ, μ) is isomorphic to the Borel–Lebesgue measure space $([0, 1], \mathfrak{B}([0, 1]), \nu)$ on the real unit interval; that is, there exists an isomorphism $i: \Sigma \to \mathfrak{B}([0, 1])$, and $\mu \circ i^{-1}$ is equivalent (in the sense of absolute continuity) to the Lebesgue measure ν. This follows from a series of lemmas of Halmos.[14]

A family of measure spaces of particular interest are those for which $\Sigma = \mathfrak{B}(\Omega)$ and $(\Omega, \mathfrak{B}(\Omega))$ is a standard Borel space. (A standard Borel space is, by definition, isomorphic to the Borel space of a Polish space. A topological space is called a Polish space if it is homeomorphic to a separable complete metric space.) This covers, for instance, all metric spaces Ω and, in particular, the spaces $\Omega = \mathbb{R}^n$ equipped with the usual Borel algebra. Now a standard Borel space is either countable or isomorphic to [0, 1] (ref. 13, Corollary A.11). The countable case will be discussed in Section 4. Here we consider the latter possibility. Hence there exists a Borel isomorphism $\varphi: \Omega \to [0, 1]$ which induces an algebraic isomorphism $i: \mathfrak{B}(\Omega) \to \mathfrak{B}(0, 1)$ via $i(B) = \varphi(B)$.

Returning to the general case, using the above isomorphism i, one may construct a positive, isometric, linear mapping

$$j: \quad M_\mu(\Omega) \equiv L^1(\Omega, \Sigma, \mu) \to M_\nu(0, 1) \equiv L^1([0, 1], \nu) \equiv L^1(0, 1)$$

for deriving the assertion of Theorem 3.1 by application of the RSS theorem. Here $M_\mu(\Omega)$ [resp. $M_\nu(0, 1)$] denotes the vector space of all finite signed measures on Σ [resp. $\mathfrak{B}(0, 1)$] which are absolutely continuous with respect to μ (resp. ν). The elements of $M_\mu(\Omega)$ are precisely the signed measures μ_f whose Radon–Nikodym derivatives are elements f of $L^1(\Omega, \Sigma, \mu)$: $\mu_f(B) = \int_B f\, d\mu$; in this sense the vector space $M_\mu(\Omega)$ is isomorphic to the space $L^1(\Omega, \Sigma, \mu)$. To every $\mu_f \in M_\mu(\Omega)$ we can assign a signed measure $\mu_f \circ i^{-1}$ on $\mathfrak{B}(0, 1)$. Any $\mu_f \in M_\mu(\Omega)$ is absolutely continuous with respect to the measure μ, and $\mu \circ i^{-1}$ is absolutely continuous with respect to ν. Consequently, the signed measures $\mu_f \circ i^{-1}$ are absolutely continuous with respect to ν. Then the Radon–Nikodym theorem states that there exists a unique function g in $L^1([0, 1], \mathfrak{B}(0, 1), \nu)$ such that

$$\mu_f \circ i^{-1}(E) = \int_E g\, d\nu, \qquad E \in \mathfrak{B}(0, 1)$$

Therefore we have $\mu_f \circ i^{-1} \in M_\nu(0, 1)$ for any $\mu_f \in M_\mu(\Omega)$. Using the equivalence of $\mu \circ i^{-1}$ and ν, we can prove the converse: $\nu_g \circ i \in M_\mu(\Omega)$ for any $\nu_g \in M_\nu(0, 1)$. Thus, it is possible to define a linear bijection $j: M_\mu(\Omega) \to M_\nu(0, 1)$, $\mu_f \mapsto \mu_f \circ i^{-1}$. That j, j^{-1} are isometries follows from the fact that both j and j^{-1}, being "trace-preserving," positive, linear mapping, are contractions. Now one proceeds precisely as in the first proof of Theorem 3.1 by application of the positive isometry j.

Now it remains to extend the result to the case of measure spaces containing atoms.

Lemma 3.6. Let (Ω, Σ, μ) be a separable, σ-finite measure space containing atoms. Then there exists a separable, σ-finite, atom-free measure space $(\mathfrak{F}, \tilde{\mu})$ and an embedding of the σ-algebra Σ into the σ-algebra $\mathfrak{F}$. Furthermore, every stochastic mapping $\Phi \in \mathrm{ST}(M_{\tilde{\mu}}(\mathfrak{F}))$ induces in a canonical way a stochastic mapping $\Phi' \in \mathrm{ST}(M_\mu(\Sigma))$.

Proof. First, we shall construct an embedding of Σ into $\mathfrak{F}$. Due to the σ-finiteness of μ, the σ-algebra Σ contains at most countably many atoms $\mathfrak{A} := \{p_i\}_{i \in N}$ $[0 < \mu(p_i) < \infty]$. [Here we assume that μ is strictly positive ($\forall A \in \Sigma,\ A \neq \varnothing: \mu(A) > 0$) because the quotient algebra $\mathfrak{B} \equiv \Sigma/\Delta_\mu$ guarantees this property.] Any measure can be decomposed uniquely into a sum $\mu = \mu_c + \mu_p$, where μ_c is continuous and μ_p is a pure point measure. The measures μ_p, resp. μ_c, vanish on $\Omega_c := \Omega$-sup $\mathfrak{A}$, resp. sup $\mathfrak{A}$. Let $\Sigma(\mathfrak{A})$ be the atomic σ-algebra generated by $\mathfrak{A}$ and $\Sigma(\Omega_c)$ the smallest σ-algebra which contains the set $\mathfrak{C} := \{E \in \Sigma \mid E \subseteq \Omega_c\}$. It is not difficult to see that $\Sigma(\Omega_c)$ contains no atoms. Thus, every $E \in \Sigma$ can be decomposed into a union $E = E_c \cup E_p$ of an element $E_c \in \Sigma(\Omega_c)$ and $E_p \in \Sigma(\mathfrak{A})$, where E_p is defined by $E_p := \bigcup \{p_i \mid p_i \in \mathfrak{A}_E\}$, $\mathfrak{A}_E = E \cap \mathfrak{A}$. The σ-additivity of μ implies

$$\mu(E) = \mu_c(E_c) + \mu_p(E_p) = \mu_c(E_c) + \sum_{i=1}^{N} \mu_{p_i}(E_p \cap p_i)$$

for $N = \#\mathfrak{A}$ and all $E \in \Sigma$. The entities μ_{p_i} denote measures on the Boolean algebras $\mathfrak{X}_i = \{p_i, \varnothing\}$ given by

$$\mu_{p_i}(x_i) := \begin{cases} \mu_p(p_i), & x_i = p_i \\ 0, & x_i = \varnothing \end{cases}$$

Let

$$\mathfrak{L} := \{(E_c, \{x_i\}_{i \in N}) \mid E_c \in \Sigma(\Omega_c), x_i \in \{p_i, \varnothing\}\}$$

$\mathfrak{L}$ becomes a Boolean σ-algebra under the lattice operations $\neg E = (\Omega_c - E_c, \{p_i - x_i^E\}_{i \in N})$, $E \cap F = (E_c \cap F_c, \{x_i^E \cap x_i^F\}_{i \in N})$, and similarly $E \cup F$. The σ-algebras $\mathfrak{L}$, Σ are isomorphic because the mapping

$$i: \quad \Sigma \to \mathfrak{L}, \quad E = E_c \cup \left(\bigcup_{i \in N} x_i \right) \mapsto (E_c, \{x_i\}_{i \in N})$$

defines a σ-isomorphism. Similarly, a σ-algebra can be defined on the set

$$\mathfrak{F} := \{(E_c, \{E_i\}_{i \in N}) \mid E_c \in \Sigma(\Omega_c), N = \#\mathfrak{A}, E_i \in \mathfrak{B}(0, 1)\}$$

$\mathfrak{B}(0, 1)$ denotes the quotient algebra of the Borel algebra with respect to the Lebesgue measure. $\mathfrak{L}$ can be embedded into $\mathfrak{F}$ in virtue of the mapping

$$\varphi: \quad \mathfrak{L} \to \mathfrak{F}, \quad E = (E_c, \{x_i\}_{i \in N}) \mapsto (E_c, \{\varphi_i(x_i)\}_{i \in N})$$

where $\varphi_i: \{p_i, \varnothing\} \to \{[0, 1], \varnothing\}$ with $\varphi_i(p_i) = [0, 1]$, $\varphi_i(\varnothing) = 0$. Now, $(\mathfrak{F}, \tilde{\mu})$ becomes a measure space by the measure

$$\tilde{\mu}: \quad \mathfrak{F} \to \mathbb{R}^+ \cup \{+\infty\}, \quad \tilde{\mu}(E) = \mu(E_c) + \sum_{i=1}^{N} \mu_L(E_i)$$

(μ_L denotes the Lebesgue measure). $\tilde{\mu}$ is a σ-additive, σ-finite measure, and the measure space $(\mathfrak{F}, \tilde{\mu})$ is separable because the measure spaces $(\Sigma(\Omega_c), \mu_c)$, $(\mathfrak{B}(0,1), \mu_L)$ have this property. Obviously, $\mathfrak{F}$ contains no atoms. Now it remains to prove that every $\Phi \in \mathrm{ST}(M_{\tilde{\mu}}(\mathfrak{F}))$ induces a $\Phi' \in \mathrm{ST}(M_\mu(\Sigma))$. Recall that $M_{\tilde{\mu}}(\mathfrak{F})$ denotes the space of finite signed measures on $(\mathfrak{F}, \tilde{\mu})$ $[M_{\tilde{\mu}}(\mathfrak{F}) \cong L^1(\mathfrak{F}, \tilde{\mu})$; $\forall v \in M_{\tilde{\mu}}(\mathfrak{F})$ $\exists f \in L^1(\mathfrak{F}, \tilde{\mu})$: $f = f_c + \sum_{i=1}^N f_i$, $v(E) = \int_{E_c} f_c \, d\mu_c + \sum_{i=1}^N \int_{E_i} f_i \, d\mu_L]$. The range of the embedding φ defines a σ-subalgebra $\tilde{\mathfrak{F}} := \varphi(\mathfrak{L})$ of $\mathfrak{F}$ (i.e., φ is a bijection of $\mathfrak{L}$ onto $\tilde{\mathfrak{F}}$). Every signed measure $v \in M_{\tilde{\mu}}(\mathfrak{F})$ induces a signed measure $\tilde{v}$ on $(\tilde{\mathfrak{F}}, \tilde{\mu})$ when restricted to the σ-subalgebra $\tilde{\mathfrak{F}}$. On the other hand, it is possible to assign to every signed measure $\tilde{v} \in M_{\tilde{\mu}}(\tilde{\mathfrak{F}})$ a function $f_{\tilde{v}}$ such that $\tilde{v}$ is of the form

$$\tilde{v}(E) = \tilde{v}_{f_{\tilde{v}}}(E) = \int_E f_c(\tilde{v}) \, d\mu_c + \sum_{i=1}^{N} \alpha_i(\tilde{v}) \int_{[0,1]} \chi_{[0,1]} \, d\mu_L$$

for all $E \in \tilde{\mathfrak{F}}$, where $\alpha_i(\tilde{v}) \in \mathbb{R}$. Obviously, $f_{\tilde{v}} \in L^1(\mathfrak{F}, \tilde{\mu})$ and $\tilde{v} = v_{f_{\tilde{v}}}|_{\tilde{\mathfrak{F}}}$. Now we are able to prove that every stochastic operator $\Phi \in \mathrm{ST}(M_{\tilde{\mu}}(\mathfrak{F}))$ induces in a natural way a stochastic operator $\tilde{\Phi} \in \mathrm{ST}(M_{\tilde{\mu}}(\tilde{\mathfrak{F}}))$ according to $\tilde{\Phi}\tilde{v} := (\Phi v_{f_{\tilde{v}}})_{|\tilde{\mathfrak{F}}}$. Obviously, $\tilde{\Phi}$ is a positive, trace-preserving mapping. Linearity follows from the equations

$$\begin{aligned}\tilde{\Phi}(\tilde{v}_1 + \tilde{v}_2) &= (\Phi v_{f_{\tilde{v}_1 + \tilde{v}_2}})|_{\tilde{\mathfrak{F}}} = (\Phi v_{f_{\tilde{v}_1} + f_{\tilde{v}_2}})_{|\tilde{\mathfrak{F}}} = (\Phi v_{f_{\tilde{v}_1}} + \Phi v_{f_{\tilde{v}_2}})|_{\tilde{\mathfrak{F}}} \\ &= (\Phi v_{f_{\tilde{v}_1}})|_{\tilde{\mathfrak{F}}} + (\Phi v_{f_{\tilde{v}_2}})|_{\tilde{\mathfrak{F}}} = \tilde{\Phi}\tilde{v}_1 + \tilde{\Phi}\tilde{v}_2\end{aligned}$$

By definition, a positive, trace-preserving, linear mapping $\tilde{\Phi}$ is a stochastic operator. Obviously, the measures $\tilde{\mu} \circ \varphi = \mu_c + \sum_{i=1}^N \mu_L \circ \varphi_i$ and $\mu = \mu_c + \sum_{i=1}^N \mu_{p_i}$ are equivalent and therefore one can apply the Radon–Nikodym theorem to introduce a bijective, positive isometry $h: M_\mu(\Sigma) \to M_{\tilde{\mu}}(\tilde{\mathfrak{F}})$, $\mu_f \mapsto \mu_f \circ \varphi^{-1}$. Thus, $\tilde{\Phi} \in \mathrm{ST}(M_{\tilde{\mu}}(\tilde{\mathfrak{F}}))$ implies that $\Phi' := h^{-1}\tilde{\Phi}h \in \mathrm{ST}(M_\mu(\Sigma))$. In this way every $\Phi \in \mathrm{ST}(M_{\tilde{\mu}}(\mathfrak{F}))$ induces a $\Phi' \in \mathrm{ST}(M_\mu(\Sigma))$. QED

This completes the proof of Theorem 3.1.

4. THE RUCH–SCHRANNER–SELIGMAN THEOREM FOR THE DISCRETE CASE

In this section we will indicate how Theorem 3.1 (or, equivalently, the original RSS theorem) applies to spaces $M(\Omega)$ of finite signed measures on a discrete measurable space (Ω, Σ). A measurable space (Ω, Σ) is called *discrete* if Σ is generated by an at most countable set of atoms; in that case Ω can be chosen to be countable or finite, $\Sigma = 2^\Omega$, and $M(\Omega) = M_v(\Omega)$, where ν denotes the counting measure on Ω. Thus we are basically dealing with $M(\Omega) \cong \mathbb{R}^n$, the goal of the original RSS theorem,[10] and $M(\Omega) \cong l^1$, the sequence space equipped with the 1-norm. The identity $M(\Omega) = M_\nu(\Omega)$ already shows that Theorem 3.1 applies to the case of discrete measurable spaces, since (Ω, Σ, ν) is a separable, σ-finite, atomic measure space. But for physical applications it is nevertheless rewarding to carry out in some detail the discretization procedure inherent in Lemma 3.6; this method will turn out to coincide with the well-known *coarse-graining* technique of statistical mechanics. As a by-product, coarse graining can be understood in terms of stochastic operators themselves. In the light of the generalized RSS theorem, this shows the striking geometric features of irreversibility.

Let (Ω, Σ, μ) be a separable measure space with μ being a σ-finite, atom-free measure. Then there exists a (finite or countable) family $\Gamma = \{\omega_n\}_{n \in I}$ in Σ of subsets of Ω with $\mu(\omega_j) < \infty$, $\bigcup_{j \in I} \omega_j = \Omega$, and $\omega_j \cap \omega_i = \emptyset$ for all $i \neq j$. Let $\Sigma(\Gamma)$ denote the smallest σ-subalgebra of Σ containing Γ [thus, $\Sigma(\Gamma) = 2^\Gamma$] and $M(\Gamma)$ [$= M_\nu(\Gamma)$] the space of finite signed measures on $\Sigma(\Gamma)$. Every function $f \in L^1(\Omega, \Sigma, \mu)$ induces a signed measure $\tilde{\mu}_f$ on $\Sigma(\Gamma)$ according to

$$\tilde{\mu}_f(E) = \sum_{\omega_i \subset E} \mu_f(\omega_i), \qquad E \in \Sigma(\Gamma) \qquad (\text{thus } \tilde{\mu}_f = \mu_f|_{\Sigma(\Gamma)})$$

On the other hand, every signed measure $\tilde{\mu} \in M(\Gamma)$ can be defined by a family $\{v_n(\tilde{\mu})\}_{n \in I}$, $v_k(\tilde{\mu}) \in \mathbb{R}$, where $\sum_{k \in I} |v_k(\tilde{\mu})| < \infty$. Define a piecewise constant function on Ω:

$$f_{\tilde{\mu}}(x) := v_k(\tilde{\mu}) \Big/ \int_{\omega_k} d\mu \qquad \text{if} \quad x \in \omega_k$$

Obviously, $f_{\tilde{\mu}} \in L^1(\Omega, \Sigma, \mu)$ and $\tilde{\mu} = \mu_{f_{\tilde{\mu}}}|_{\Sigma(\Gamma)}$. Then one easily realizes that a stochastic operator $\Phi \in \mathrm{ST}(M_\mu(\Omega))$ reduces in a natural way to a stochastic operator $\tilde{\Phi} \in \mathrm{ST}(M(\Gamma))$ according to the following:

$$\tilde{\Phi}\tilde{\mu} := (\Phi\mu_{f_{\tilde{\mu}}})^{\sim} = (\Phi\mu_{f_{\tilde{\mu}}})|_{\Sigma(\Gamma)}$$

Interpreting Ω as the phase space of a physical system, the reduction of the measurable space (Ω, Σ) to $(\Gamma, \Sigma(\Gamma))$ corresponds to a partitioning into *finite* cells. The identification of the elements from $M(\Gamma)$ with piecewise constant functions in $L^1(\Omega, \Sigma, \mu)$ can be represented by means of a stochastic mapping Ψ on $M_\mu(\Omega)$ [resp. on $L^1(\Omega, \Sigma, \mu)$]: for $\mu_f \in M_\mu(\Omega)$, $f \in L^1(\Omega, \Sigma, \mu)$, put

$$\Psi\mu_f(E) := \sum_{i \in I} \mu(E \cap \omega_i)\, \mu_f(\omega_i)/\mu(\omega_i)$$

Thus

$$\Psi f(x) = \mu_f(\omega_i)/\mu(\omega_i) \qquad \text{if} \quad x \in \omega_i$$

Now the range of Ψ in $M_\mu(\Omega)$ is isomorphic to $M(\Gamma)$. Let $\Theta: M(\Gamma) \to \Psi M_\mu(\Omega)$, Θ^{-1} denote the corresponding positive, trace-preserving (stochastic) isometries. Then the above canonical relation between stochastic operators on $M_\mu(\Omega)$ and $M(\Gamma)$ can be summarized as $\tilde{\Phi} := \Theta^{-1} \circ \Psi \circ \Phi \circ \Theta$. Moreover, interpreting Φ as a dynamical mapping on $M_\mu(\Omega)$, then $\tilde{\Phi}$ is nothing but the canonically associated *coarse-grained* "image" dynamics on $M(\Gamma)$. Being the composition of several stochastic operators of which at least one (Ψ) is not an isometry, it is clear that $\tilde{\Phi}$ is itself no isometry. On the contrary, it can be considered to be "more mixing" than the original dynamics Φ. In this way we have interpreted irreversibility introduced by means of coarse graining in terms of stochastic operations.

5. CONCLUSION: POSSIBLE FIELDS OF APPLICATIONS

In the introduction we mentioned a number of research areas where the notions of mixing character and mixing distance provided new insights into natural phenomena as well as mathematical structures. Having extended the scope of the Ruch–Schranner–Seligman theorem to a fairly large class of L^1 spaces, we shall conclude this paper by indicating by means of two examples how the Principle of Decreasing Mixing Distance governs the statistical description of dynamical evolution processes.

First we briefly sketch the general features of population dynamics as it occurs in biological ecosystems, chemical reactions, laser physics, or molecular evolution.[9,15] This yields an example for the finite discrete case of the RSS theorem. The phenomena mentioned above admit various approaches for a quantitative description going beyond the basic level of *reaction equations*. One may, for instance, formulate *rate equations* for the occupation numbers corresponding to a finite number of *states* (excitation states, species). In certain circumstances such reactions are Markov processes governed by a stochastic matrix of transition probabilities. But in general one is dealing with nonlinear phenomena, that is, nonlinear equations for the respective phase space trajectories. Then the RSS theorem suggests adopting a statistical level of description, i.e., ascribing *occupation probabilities* to the occupation numbers and trying to write down evolution equations for the probability distributions. It turns out that even when one starts with a nonlinear (stochastic) dynamical equation, the resulting induced statistical evolution is a *linear* process. This phenomenon is well known in the context of the second example discussed in Section 4, phase space statistical mechanics.

The probabilistic representation of the *nonlinear* Hamiltonian dynamics yields the *linear* Liouville equation, the solution of which is given by a group Φ_t of *isometric* stochastic operators; thus, in that case the mixing distance for arbitrary pairs of probability measures is time invariant. Irreversibility occurs only on the macroscopic level of a simplified description in terms of coarse graining: in that case one introduces a phase-space cell partition and formulates evolution equations for the time development of probability measures defined on the resulting *discrete* event space. The resulting irreversible dynamics is represented as a stochastic semigroup $\tilde{\Phi}_t$. This is an instance of the RSS theorem applied to *countable* discrete measurable spaces.

To reemphasize, a most remarkable aspect of dynamical processes brought into perspective by considering the Principle of Decreasing Mixing Distance is that in principle any *nonlinear* evolution equation can be represented by a family of *linear* stochastic operators acting on the probability distributions defined on the original *phase space*. This may turn out to be the starting point for practical applications of the Principle in the description of general features of complex phenomena of organization. Its *fundamental* importance, however, is rooted in its geometrical characterization of irreversibility.

ACKNOWLEDGMENTS

It is a pleasure to thank E. Ruch for allowing us to quote ideas laid down in his work of ref. 8 prior to its publication. Partial support by the DFG, Bonn, is gratefully acknowledged. Finally, thanks are due to W. Stulpe for useful comments on the manuscript.

REFERENCES

1. E. Ruch and A. Schönhofer, Theorie der Chiralitätsfunktionen, *Theor. Chim. Acta* (Berlin) **19**:225–287 (1970).
2. E. Ruch, The diagram lattice as a structural principle, *Theor. Chim. Acta* (Berlin) **38**:167–183 (1975).
3. E. Ruch and I. Gutman, The branching extent of graphs, *J. Combinatorics, Information Syst. Sci.* **4**:285–295 (1975).
4. E. Ruch and A. Mead, The principle of increasing mixing character and some of its consequences, *Theor. Chim. Acta* (Berlin) **41**:95–117 (1976).
5. E. Ruch and B. Lesche, Information extent and information distance, *J. Chem. Phys.* **69**:393–401 (1978).
6. F. Schlögl, Mixing distance and stability of steady states in statistical nonlinear thermodynamics, *Z. Physik B* **25**:411–421 (1976).

7. F. Schlögl and C. A. Mead, Macroscopic aspects of mixing distance in nonequilibrium thermodynamics, *Ann. Phys.* (N.Y.) **155**:172–190 (1978).
8. E. Ruch, The direction distance, preprint (1990).
9. H. Haken, *Synergetics* (Springer, Berlin, 1978).
10. E. Ruch, R. Schranner, and T. H. Seligman, Generalization of a theorem by Hardy, Littlewood, and Polya, *J. Math. Anal. Appl.* **76**:222–229 (1980).
11. P. Busch and E. Ruch, The measure cone—Irreversibility as a geometrical phenomenon, Preprint (1990).
12. P. M. Alberti and A. Uhlmann, *Math. Nachr.* **97**:279–295 (1980).
13. M. Takesaki, *Theory of Operator Algebras* (Springer, Berlin, 1979).
14. P. R. Halmos, *Measure Theory* (Van Nostrand, New York, 1950).
15. A. Dress, Populationsdynamik auf Sequenzräumen, in *Ordnung aus dem Chaos*, B.-O. Küppers, ed. (Piper, Munich, 1987).