Commun. Math. Phys. 105, 13–34 (1986)

Communications in Mathematical Physics

On The Relative Entropy

Matthew J. Donald
18 Holland Street, Cambridge CB4 3DL, Great Britain

Abstract. For $\mathscr{A}$ any subset of $\mathscr{B}(\mathscr{H})$ (the bounded operators on a Hilbert space) containing the unit, and σ and ρ restrictions of states on $\mathscr{B}(\mathscr{H})$ to $\mathscr{A}$, $\text{ent}_{\mathscr{A}}(\sigma|\rho)$—the entropy of σ relative to ρ given the information in $\mathscr{A}$—is defined and given an axiomatic characterisation. It is compared with $\text{ent}^S_{\mathscr{A}}(\sigma|\rho)$—the relative entropy introduced by Umegaki and generalised by various authors—which is defined only for $\mathscr{A}$ an algebra. It is proved that ent and ent^S agree on pairs of normal states on an injective von Neumann algebra. It is also proved that ent always has all the most important properties known for ent^S: monotonicity, concavity, w^* upper semicontinuity, etc.

1. Introduction

Given states σ and ρ on a von Neumann algebra $\mathscr{A}$, the entropy of σ relative to ρ, written $\text{ent}_{\mathscr{A}}(\sigma|\rho)$, is a measure of how easy it is to distinguish the state ρ from the state σ. As such it has, since first introduced by Umegaki [1], found application in quantum statistical mechanics [2, 3 (Sect. 6.2.3, pp 269–289), 4], quantum information theory [1, 5], the foundations of quantum mechanics [6], and the theory of von Neumann algebras [7, 8]. I shall give a brief sketch below of how I see the role of relative entropy in the foundations of quantum theory, as this is my own motivation for studying the subject. If my view of these matters is correct then the relative entropy is of fundamental importance to physics.

As a mathematical object the relative entropy is fascinating. It has been given three distinct but equivalent definitions: that of Araki [9, 10] who uses Tomita–Takesaki theory, that of Pusz and Woronowicz [11] who use their "functional calculus for sesquilinear forms" [12], and that of Uhlmann [13] who uses interpolation theory. The entropy defined by Araki and Uhlmann will be denoted by $\text{ent}^S_{\mathscr{A}}(\sigma|\rho)$ throughout this paper. For the equivalence of their definitions see [8]. The entropy of Pusz and Woronowicz will be denoted by $\text{ent}^{PW}_{\mathscr{A}}(\sigma|\rho)$. Its equivalence to $\text{ent}^S_{\mathscr{A}}(\sigma|\rho)$ will be proved in Sect. 4 of this paper.

The purpose of this paper is to give yet another definition. This new definition has several advantages. It is given by means of a set of axioms and is conceptually and mathematically simpler than the previous definitions. It gives a characterisation of the relative entropy which allows for a heuristic interpretation. This is significant,