Hokkaido Mathematical Journal Vol. 18 (1989) p. 45～77

The Finsler geometry of certain covering groups of operator groups

C. J. ATKIN

(Received September 7, 1987, Revised December 14, 1987)

In a previous paper [1], I have investigated the Finsler geometry, in a natural Finsler structure, of both the full and the Fredholm unitary, orthogonal, and symplectic groups of a Hilbert space E. In this sequel I shall discuss similar questions for the associated non-trivial covering groups, the Fredholm spinor group $\boldsymbol{SpinC}(E)$ (the universal cover of the special Fredholm orthogonal group $\boldsymbol{SOC}(E)$) and the coverings of the Fredholm unitary group $\boldsymbol{UC}(E)$.

There is no evident relation between the global Finsler geometry of a covering group and that of the original group, unless one remains sufficiently close to the identity. Thus some of my results, whilst not deep, are perhaps a little unexpected. The most striking are these:

(i) The universal cover of $\boldsymbol{UC}(E)$ is bounded when, and only when, E is of infinite dimension (see (6.7), (7.5)(b)). This seems contrary to intuition, but the intuition is perhaps founded on the fact that the n-fold cover of the finite-dimensional group $\boldsymbol{U}(n)$ is the product of $\boldsymbol{SU}(n)$ with the circle group, and its universal cover is therefore just $\boldsymbol{SU}(n)\times\boldsymbol{R}$. When E is infinite-dimensional, $\boldsymbol{UC}(E)$ admits no determinant homomorphism, and its universal cover cannot be so explicitly represented.

(ii) In the same universal cover of $\boldsymbol{UC}(E)$, the exponential map is onto (see (6.8)(a), (7.6)(a)). However, if E is separably infinite-dimensional, and only then, there exist pairs of points in this group which cannot be joined by a minimising path (see (6.9)). The curious dependence on dimension arises because, when E is non-separable, there is great freedom to construct paths in $\boldsymbol{UC}(E)$, and, when E is finite-dimensional, there are relatively few points to consider.

(iii) The properties of $\boldsymbol{SpinC}(E)$ are largely independent of the dimension of E, provided that it is greater than 2. However, a point in $\boldsymbol{SpinC}(E)$ may be joined to the identity either by exactly one minimising one-parameter subgroup, if it is close enough; or by infinitely many; or, surprisingly, by exactly two (see (5.8)(c)).

Except for minor changes of notation, the arguments here follow, and occasionally presuppose, those of [1]. In § 1, the topics and methods are introduced. In §§ 2, 3, spectral theory is used to derive a homotopy classification (the 'twist') of loops in $\boldsymbol{SOC}(E)$; an intuitive interpretation is given at (3.12), and makes the computation of the twist easy wherever it is needed subsequently. The analogous concept for $\boldsymbol{UC}(E)$ is the 'degree', presented in § 4. Then § 5 studies the Finsler geometry of $\boldsymbol{SpinC}(E)$ and 6 that of the universal cover of $\boldsymbol{UC}(E)$ for infinite-dimensional E. Finally, § 7 discusses (rather cursorily) the finite coverings of $\boldsymbol{UC}(E)$ and the case of finite-dimensional E.

Throughout the paper, minimising paths are understood to be uniformly parametrised. As in [1], no attempt is made to study the qualitative behaviour of geodesics, but it is worth noting that, whenever I say there are infinitely many minimising geodesics, there is clearly a continuous family of them.

§ 1. Preliminaries.

(1.1) Let E be a complex Hilbert space (usually of infinite dimension), furnished with a real structure J-that is, in the terminology of [1], an involutive conjugate-isometry of E, which in effect represents complex conjugation. The Fredholm unitary group $\boldsymbol{UC}(E)$ of E is the group of all unitary operators T in E such that $T-I$ is a compact operator (where I is the identity in E). The Fredholm orthogonal group of E, here denoted $\boldsymbol{OC}(E)$, is the subgroup of $\boldsymbol{UC}(E)$ consisting of operators which commute with J. Both $\boldsymbol{UC}(E)$ and $\boldsymbol{OC}(E)$ are Banach Lie groups, and are furnished with left- and right-invariant Finsler structures as in §§ 1, 2 of [1].

Most of the arguments in this paper do not require that E be infinite-dimensional; see (7.4) and (3.12) for indications.

(1.2) There is a homotopy equivalence between $\boldsymbol{OC}(E)$ and $\boldsymbol{O}=\lim_{n\to\infty} O(n)$. For a proof which does not require separability of E, see [2]. It follows that $\boldsymbol{OC}(E)$ has two components; the principal component is the special Fredholm orthogonal group $\boldsymbol{SOC}(E)$, and an element T of $\boldsymbol{OC}(E)$ belongs to $\boldsymbol{SOC}(E)$ if and only if $\ker(T+I)$ is of even dimension (see § 5 of [1]). The fundamental group of $\boldsymbol{SOC}(E)$ is isomorphic to $\boldsymbol{Z}_2$, and its universal cover is the Fredholm spinor group $\boldsymbol{SpinC}(E)$. Since the covering map $\boldsymbol{\pi}$: $\boldsymbol{SpinC}(E)\longrightarrow \boldsymbol{SOC}(E)$ is a local isomorphism, $\boldsymbol{SpinC}(E)$ inherits both a Banach Lie group structure and a unique Finsler structure such that the tangent map of $\boldsymbol{\pi}$ is an isometry of each tangent space. This Finsler structure, which is clearly both left- and right-invariant, is the one assumed henceforth.

As in [1], the Finsler length of a rectifiable path will be denoted by ℓ, and the induced metric by d, irrespective of the group under consideration. Write also $B(T\,;\varepsilon)$ for the set of elements of the group whose distance from the given element T is less than ε.

(1.3) No satisfactory explicit construction of $\boldsymbol{SpinC}(E)$ has to my knowledge been discovered (for some closely related groups, see [3] or [5]), and I shall therefore base my investigation on the standard construction of covering spaces (see, for instance, [4], pp. 253 et seqq.)

Let $P(\boldsymbol{SOC}(E))$ denote the class of rectifiable paths $p:[0,1]\longrightarrow \boldsymbol{SOC}(E)$ for which $p(0)=e$, the identity element of $\boldsymbol{SOC}(E)$, and let

$$\tau : P(\boldsymbol{SOC}(E))\longrightarrow \boldsymbol{SOC}(E) : p\longmapsto p(1).$$

In the same way, one has the terminal evaluation

$$\tilde{\tau} : P(\boldsymbol{SpinC}(E))\longrightarrow \boldsymbol{SpinC}(E).$$

Composition with $\boldsymbol{\pi}$ defines a bijection

$$\boldsymbol{\pi}_* : P(\boldsymbol{SpinC}(E))\longrightarrow P(\boldsymbol{SOC}(E))$$

(with lifting as its inverse), which is length-preserving, by (1.2). Certainly $\boldsymbol{\pi}\tilde{\tau}=\tau\boldsymbol{\pi}_*$, and $\tilde{\tau}(p_1)=\tilde{\tau}(p_2)$ if and only if $\boldsymbol{\pi}_*p_1$ and $\boldsymbol{\pi}_*p_2$ are homotopic modulo $\{0,1\}$, that is, homotopic with fixed end-points.

(1.4) The points of $\boldsymbol{SpinC}(E)$ will be represented as homotopy equivalence classes modulo $\{0,1\}$ of elements of $P(\boldsymbol{SOC}(E))$. The projection $\boldsymbol{\pi}$ corresponds to the quotient mapping induced from τ; and the rectifiable paths from the identity $\tilde{e}$ of $\boldsymbol{SpinC}(E)$ to a given point $\tilde{x}\in\boldsymbol{SpinC}(E)$ are the lifts of paths in the homotopy class representing $\tilde{x}$. Thus the Finsler metric d in $\boldsymbol{SpinC}(E)$ is characterised by the assertion that $d(\tilde{e},\tilde{x})$ is the infimum of the lengths of paths in $P(\boldsymbol{SOC}(E))$ representing $\tilde{x}$. There is a minimising path from $\tilde{e}$ to $\tilde{x}$ in $\boldsymbol{SpinC}(E)$ if and only if that infimum is attained, and there is a minimising geodesic from $\tilde{e}$ to $\tilde{x}$ if and only if it is attained by a segment of a one-parameter subgroup in $\boldsymbol{SOC}(E)$. In short, I must study here the minimisation of length over homotopy classes, instead of over all paths as in [1].

It follows that the core of this paper must be an investigation of the homotopy relation for paths in terms of spectral theory (which decides distances—see [1]). As far as I know, this has not been done before. The method is to associate to certain paths a sort of holonomy transformation, and it may be developed much further, but it would be inappropriate to discuss that here.

(1.5) The Fredholm unitary group $\boldsymbol{UC}(E)$ of E consists of all the complex-linear isometries T of E for which $T-I$ is compact. Its homotopy type (see [2]) is that of $\boldsymbol{U}=\lim_{n\to\infty} U(n)$, so that it is connected, with fundamental group isomorphic to $\boldsymbol{Z}$. The remarks (1.3) and (1.4) have obvious extensions to $\boldsymbol{UC}(E)$ and its covering groups.

§ 2. Coupled subspaces.

Suppose in what follows that F_i is a closed subspace of E, for $i=1, 2, 3$, and that P_i is the orthogonal projection on F_i.

(2.1) Say that F_1 and F_2 (or P_1 and P_2) are coupled if $P_1|F_2 : F_2 \longrightarrow F_1$ and $P_2|F_1 : F_1 \longrightarrow F_2$ are both isomorphisms. Thus F_1 and F_2 must have the same dimension. When F_1 and F_2 are finite-dimensional, as they usually will be in this paper, F_1 and F_2 will be coupled if and only if $F_1 \cap F_2^{\perp}=0=F_1^{\perp} \cap F_2$. Indeed, $F_1 \cap F_2^{\perp}=0$ means that $P_2|F_1$ is one-one, whilst $F_1^{\perp} \cap F_2=0$ means that $P_2|F_1$ is surjective.

(2.2) Next, suppose that each F_i is a real subspace of E—that is, invariant under J. This is equivalent to saying P_i is a real operator (commuting with J). Now, if F_1 and F_2 are also coupled and finite-dimensional, they may be oriented, and any orientation ω of F_1 induces, via the real isomorphism $P_2|F_1 : F_1 \longrightarrow F_2$, an orientation $P_2\omega$ of F_2. This notation is unambiguous, since ω orients the specific subspace F_1.

(2.3) Let F be any nonzero complex finite-dimensional subspace of E. Then $F+J(F)$ is a real finite-dimensional subspace; and, when $F \cap J(F)=0$, there is a canonical orientation for $F+J(F)$. Indeed, take any ordered basis (over $\boldsymbol{C}$) for $F : x_1, x_2, \cdots, x_m$. The associated ordered real basis for $F+J(F)$ given by x_1+Jx_1, $i(x_1-Jx_1)$, x_2+Jx_2, $i(x_2-Jx_2)$, $\cdots$ determines the required orientation; it is easily verified that the initial choice of basis in F does not influence the result, and even that, when F and $J(F)$ are interchanged, the orientation of $F+J(F)$ will be reversed only if F is of odd dimension.

(2.4) As in (3.3) of [1], δ denotes the metric on the complex unit circle S: $\delta(\exp(i\theta), \exp(i\varphi))=\min\{|\theta-\varphi+2n\pi| : n \in \boldsymbol{Z}\}$;

and $\boldsymbol{P}_T$ is the spectral measure associated to the bounded normal operator T. For the moment, let $\mathfrak{G}$ denote any of the groups considered in [1] (see (1.6), (2.2) of that paper), with Finsler metric d. Given $\mu \in S$, $\beta \geq 0$, and $U \in \mathfrak{G}$, set

$$\boldsymbol{Q}_U(\mu, \beta) = \boldsymbol{P}_U(\{z \in S : \delta(z, \mu) \leq \beta\}) \text{ and}$$
$$K_U(\mu, \beta) = \boldsymbol{Q}_U(\mu, \beta)E.$$

(2.5) PROPOSITION. *Let $T \in \mathfrak{G}$, $\mu \in S$, $\beta \geq 0$, $6\varepsilon \in (0, \pi]$. Suppose that*

$$\{z \in S : \delta(z, \mu) \leq \beta + 6\varepsilon\} \cap \sigma(T) = \{z \in S : \delta(z, \mu) \leq \beta\} \cap \sigma(T). \tag{1}$$

Then (a) *for $U \in \mathfrak{G}$, $\boldsymbol{Q}_U(\mu, \rho)$ is independent of ρ for*

$$\beta + d(T, U) \leq \rho \leq \beta + 6\varepsilon - d(T, U);$$

(b) *for $U \in B(T; 3\varepsilon)$, $\boldsymbol{Q}_U(\mu, \beta + d(T, U))$ is a C^ω operator-valued function of U with respect to the operator-norm;*

(c) *if $U, V \in \mathfrak{G}$ and*

$$\min(d(T, U), d(T, V)) + d(U, V) \leq 3\varepsilon,$$
$$d(T, U) + d(U, V) + d(T, V) \leq 6\varepsilon,$$

then $\boldsymbol{Q}_U(\mu, \beta + d(T, U))$ and $\boldsymbol{Q}_V(\mu, \beta + d(T, V))$ are coupled.

PROOF. When $d(T, U) < 6\varepsilon \leq \pi$, (7.2) of [1] gives a rectifiable path p in $\mathfrak{G}$ with $p(0) = T$, $p(1) = U$, $\ell(p) = d(T, U)$. Thus, by (4.7) of [1], for any $\rho \geq 0$

$$K_U(\mu, \rho + d(T, U))^\perp \cap K_T(\mu, \rho) = 0. \tag{2}$$

By hypothesis (1), however,

$$K_T(\mu, \beta) = K_T(\mu, \rho) = K_T(\mu, \beta + 6\varepsilon) \tag{3}$$

for $\beta \leq \rho \leq \beta + 6\varepsilon$. Applying (4.7) of [1] again,

$$\begin{aligned} &K_U(\mu, \rho + d(T, U)) \cap K_T(\mu, \rho + 2d(T, U))^\perp \\ &\quad = K_U(\mu, \rho + d(T, U)) \cap K_T(\mu, \beta)^\perp = 0 \end{aligned} \tag{4}$$

when $\beta \leq \rho \leq \beta + 6\varepsilon - 2d(T, U)$. But (2) and (4) prove that $\boldsymbol{Q}_U(\mu, \rho + d(T, U))$ and $\boldsymbol{Q}_T(\mu, \rho)$ are coupled, for

$$\beta \leq \rho \leq \beta + 6\varepsilon - 2d(T, U). \tag{5}$$

The projection $\boldsymbol{Q}_U(\mu, \rho + d(T, U))$ increases with ρ:

$$K_U(\mu, \rho_1 + d(T, U)) \subseteq K_U(\mu, \rho_2 + d(T, U)) \text{ when } \rho_1 \leq \rho_2.$$

Suppose $\beta \leq \rho_1 \leq \rho_2 \leq \beta + 6\varepsilon - 2d(T, U)$, and

$$x \in K_U(\mu, \rho_2 + d(T, U)) \cap K_U(\mu, \rho_1 + d(T, U))^\perp. \tag{6}$$

From (5), there is a unique $y \in K_T(\mu, \rho_2)$ such that

$$x = \boldsymbol{Q}_U(\mu, \rho_2 + d(T, U))y;$$

consequently

$$\begin{aligned}0&=\langle \boldsymbol{Q}_U(\mu, \rho_2+d(T, U))y,\ \boldsymbol{Q}_U(\mu, \rho_1+d(T, U))y\rangle\\ &=\langle \boldsymbol{Q}_U(\mu, \rho_1+d(T, U))\boldsymbol{Q}_U(\mu, \rho_2+d(T, U))y,\ \boldsymbol{Q}_U(\mu, \rho_1+d(T, U))y\rangle\\ &=\langle \boldsymbol{Q}_U(\mu, \rho_1+d(T, U))y,\ \boldsymbol{Q}_U(\mu, \rho_1+d(T, U))y\rangle.\end{aligned}$$

Hence $\boldsymbol{Q}_U(\mu, \rho_1+d(T, U))y=0$ and

$$y\in K_T(\mu, \rho_2)\cap K_U(\mu, \rho_1+d(T, U))^{\perp}.$$

In view of (3), this means $y\in K_T(\mu, \rho_1)\cap K_U(\mu, \rho_1+d(T, U))^{\perp}$, and therefore, by (2), $y=0$. Thus $x=0$. Recalling (6),

$$\boldsymbol{Q}_U(\mu, \rho_2+d(T, U))=\boldsymbol{Q}_U(\mu, \rho_1+d(T, U))$$

for $\beta\leq\rho_1\leq\rho_2\leq\beta+6\varepsilon-2d(T, U)$; this in effect is (a). It follows that, for $U\in B(T\,;3\varepsilon)$,

$$\begin{aligned}&\{z\in S: \delta(z, \mu)\leq\beta+d(T, U)\}\cap\sigma(U)\\ &\quad=\{z\in S: \delta(z, \mu)<\beta+6\varepsilon-d(T, U)\}\cap\sigma(U).\end{aligned}\tag{7}$$

Choose $\gamma>0$ so that

$$\{z\in S: \delta(z, \mu)\leq\beta+3\varepsilon\}=\{z\in S: |z-\mu|\leq\gamma\}.$$

Then, by (7), $\{z\in C: |z-\mu|=\gamma\}\cap\sigma(U)=\phi$; this holds for any $U\in B(T\,; 3\varepsilon)$, so it follows that for such U

$$\boldsymbol{Q}_U(\mu, \beta+d(T, U))=(2\pi i)^{-1}\int_{|z-\mu|=\gamma}(z\,I-U)^{-1}dz,$$

which is a C^{ω} function of U. This is (b).

In turn, (7) means that, when $d(T, U)<3\varepsilon$, (5) may be applied with U in place of T, $\beta+d(T, U)$ in place of β, and $6\varepsilon-2d(T, U)$ in place of 6ε. One infers that, given $V\in\mathfrak{S}$, $\boldsymbol{Q}_U(\mu, \tau)$ and $\boldsymbol{Q}_V(\mu, \tau+d(U, V))$ are coupled when

$$\beta+d(T, U)\leq\tau\leq\beta+d(T, U)+6\varepsilon-2d(T, U)-2d(U, V).$$

In particular, $\boldsymbol{Q}_U(\mu, \beta+d(T, U))$ and $\boldsymbol{Q}_V(\mu, \beta+d(T, U)+d(U, V))$ are coupled provided that

$$2d(T, U)+2d(U, V)\leq 6\varepsilon.\tag{8}$$

On the other hand, $\boldsymbol{Q}_V(\mu, \rho)$ is independent of ρ for

$$\beta+d(T, V)\leq\rho\leq\beta+6\varepsilon-d(T, V), \text{ by (a); so}$$
$$\boldsymbol{Q}_V(\mu, \beta+\mathrm{d}(T, V))=\boldsymbol{Q}_V(\mu, \beta+d(T, U)+d(U, V))$$

provided that

$$d(T, U)+d(U, V)+d(T, V)\leq 6\varepsilon. \tag{9}$$

The two conditions (8), (9) are in effect those stated in (c), since U and V appear symmetrically in the conclusion.

(2.6) PROPOSITION. *Suppose in* (2.5) *that* $\mathfrak{G}=\boldsymbol{SOC}(E)$, $\mu=-1$.

Let $V, U, W\in B(T;\varepsilon)$, *where* $\beta+\varepsilon<\pi$. *Then* $\boldsymbol{Q}_U(-1, \beta+d(T, U))$, $\boldsymbol{Q}_V(-1, \beta+d(T, V))$, $\boldsymbol{Q}_W(-1, \beta+d(T, W))$ *are real operators of finite rank. If* ω *is an orientation of* $K_U(-1, \beta+d(T, U))$, *then*

$$\begin{aligned}\boldsymbol{Q}_V(-1, \beta+d(T, V))\boldsymbol{Q}_W(-1, \beta+d(T, W))\omega \\ =\boldsymbol{Q}_V(-1, \beta+d(T, V))\omega.\end{aligned}$$

PROOF. Since $\beta+\varepsilon<\pi$, $U|K_U(-1, \beta+d(T, U))$ has spectrum bounded away from 1 in S, and therefore consisting of finitely many eigenvalues of finite multiplicity; so $K_U(-1, \beta+d(T, U))$ is of finite dimension, and similarly for V and W. The reality of the subspaces and of the associated projections is obvious.

Now let $q:[0, 1]\longrightarrow B(T;\varepsilon)$ be a continuous path such that $q(0)=V$, $q(1)=W$. Such paths exist, for instance as the concatenation of rectifiable paths between T and V, W respectively, whose lengths approximate $d(T, V)$, $d(T, W)$ sufficiently closely. By (2.5)(b), $\boldsymbol{Q}_{q(t)}(-1, \beta+d(q(t), T))$ is continuous in t in the norm-topology. Ergo,

$$r(t)=\boldsymbol{Q}_V(-1, \beta+d(T, V))\boldsymbol{Q}_{q(t)}(-1, \beta+d(q(t), T))$$

is continuous in t, and by (2.5)(c) it constitutes, for each $t\in[0, 1]$, an isomorphism of $K_U(-1, \beta+d(T, U))$ with $K_V(-1, \beta+d(T, V))$. Thus $r(0)\omega=r(1)\omega$, and this is precisely the assertion of the Proposition.

(2.7) COROLLARY. *With the hypotheses of* (2.6),

(a) $\boldsymbol{Q}_U(-1, \beta+d(T, U))\boldsymbol{Q}_V(-1, \beta+d(T, V))\omega=\omega$,

(b) $\boldsymbol{Q}_U(-1, \beta+d(T, U))\boldsymbol{Q}_V(-1, \beta+d(T, V))\cdot\boldsymbol{Q}_W(-1, \beta+d(T, W))\omega=\omega$.

PROOF. For (a), read U for V, and V for W, in (2.6). For (b), apply $\boldsymbol{Q}_U(-1, \beta+d(T, U))$ on both sides in (2.6), and use (a). (In fact (a) may be proved far more generally than this.)

(2.8) Again let $\mathfrak{G}=\boldsymbol{SOC}(E)$, $U\in\mathfrak{G}$, $0\leq\eta\leq\theta<\pi$. Then

$$\boldsymbol{P}_U(\{z\in S:\eta<\delta(-1, z)\leq\theta\})=\boldsymbol{Q}_U(-1, \theta)-\boldsymbol{Q}_U(-1, \eta)$$

is real and of finite rank (as in (2.6)). Its image $K_U(-1, \theta)\cap K_U(-1, \eta)^{\perp}$

is the orthogonal direct sum of

$$F(\eta, \theta, U)=\boldsymbol{P}_U(\{z\in S : \eta<\delta(-1, z)\leq\theta \quad and \quad \mathrm{Im}(z)>0\})E,$$
$$J(F(\eta, \theta, U))=\boldsymbol{P}_U(\{z\in S : \eta<\delta(-1, z)\leq\theta \quad and \quad \mathrm{Im}(z)<0\})E.$$

As such, it is either zero, or has a specific orientation, which I call the standard orientation, determined as in (2.3). (If $F(\eta, \theta, U)$ is of odd dimension, it is essential to observe that it involves the inequality $\mathrm{Im}(z)>0$.) Now, say that orientations on $K_U(-1, \eta)$ and on $K_U(-1, \theta)$ are coherent if either $K_U(-1, \theta)=K_U(-1, \eta)$ and the orientations are the same, or the orientation on $K_U(-1, \theta)$ is the product of that on $K_U(-1, \eta)$ with the standard orientation on $K_U(-1, \theta)\cap K_U(-1, \eta)^{\perp}$. Coherence is easily seen to be an equivalence relation.

(2.9) PROPOSITION. *Suppose* $T, U\in\boldsymbol{SOC}(E)$, $0\leq\alpha<\beta$: *let* $\varepsilon>0$ *be such that* $\beta+3\varepsilon<\pi$, *and assume that*

$$\{z\in S : \delta(z, -1)\leq\beta+6\varepsilon\}\cap\sigma(T)=\{z\in S : \delta(z, -1)\leq\beta\}\cap\sigma(T), \quad (1)$$
$$\{z\in S : \delta(z, -1)\leq\alpha+6\varepsilon\}\cap\sigma(T)=\{z\in S : \delta(z, -1)\leq\alpha\}\cap\sigma(T). \quad (2)$$

Then, if $d(T, U)<3\varepsilon$ *and* ω_1, ω_2 *are coherent orientations of* $K_T(-1, \alpha)$ *and of* $K_T(-1, \beta)$ *respectively*, $\boldsymbol{Q}_U(-1, \alpha+d(T, U))\omega_1$ *and* $\boldsymbol{Q}_U(-1, \beta+d(T, U))\omega_2$ *are coherent orientations of* $K_U(-1, \alpha+d(T, U))$ *and of* $K_U(-1, \beta+d(T, U))$.

(See (2.2) and (2.5)(c) for the admissibility of these claims.)

PROOF. If $\alpha+6\varepsilon\geq\beta$, it follows from (1) and (2) that

$$\boldsymbol{Q}_T(-1, \alpha)=\boldsymbol{Q}_T(-1, \beta)=\boldsymbol{Q}_T(-1, \beta+6\varepsilon).$$

Apply (2.5)(a) with $\beta-\alpha+6\varepsilon$ in place of 6ε, noting that $\alpha+\beta-\alpha+6\varepsilon-d(T, U)>\beta+d(T, U)$; thus

$$\boldsymbol{Q}_U(-1, \alpha+d(T, U))=\boldsymbol{Q}_U(-1, \beta+d(T, U)).$$

So in this case there is nothing to prove, and we may assume that $\alpha+6\varepsilon<\beta$.

As $\sigma(T)$ is compact, there exists $\alpha_1\in(\alpha, \beta-6\varepsilon)$ such that

$$\{z\in S : \delta(z, -1)\leq\alpha_1+6\varepsilon\}\cap\sigma(T)=\{z\in S : \delta(z, -1)\leq\alpha\}\cap\sigma(T). \quad (3)$$

Take $\gamma=(\alpha_1+6\varepsilon+\beta)/2$, $\nu=\exp\{i(\pi-\gamma)\}$, $\zeta=(\beta-\alpha_1-6\varepsilon)/2$. Thus

$$\sigma(T)=\sigma(T)\cap(\{z\in S : \delta(z, -1)\leq\alpha\}\cup\{z\in S : \delta(z, \nu)\leq\zeta\}$$
$$\cup\{z\in S : \delta(z, \bar{\nu})\leq\zeta\}\cup\{z\in S : \delta(z, 1)\leq\pi-\beta-6\varepsilon\}), \quad (4)$$

where each of the sets in braces is δ-distant at least 6ε from each of the others. Recalling (7) of (2.5), one sees that, since $d(T, U)<3\varepsilon$,

$$\begin{aligned}\sigma(U)&=\sigma(U)\cap(\{z\in S:\delta(z,-1)\leq\alpha+d(T,U)\}\\&\qquad\cup\{z\in S:\delta(z,-1)\geq\alpha+3\varepsilon\})\\&=\sigma(U)\cap(\{z\in S:\delta(z,\nu)\leq\zeta+d(T,U)\}\\&\qquad\cup\{z\in S:\delta(z,\nu)\geq\zeta+3\varepsilon\}),\end{aligned}$$

and likewise for the other centres $\bar{\nu}$ and $+1$, with corresponding radii ζ and $\pi-\beta-6\varepsilon$. Hence I may express $\sigma(U)\cap\{z\in S:\delta(z,-1)\leq\beta+d(T,U)\}$ as the disjoint union

$$\begin{aligned}&\sigma(U)\cap(\{z\in S:\delta(z,-1)\leq\alpha+d(T,U)\}\cup\{z\in S:\delta(z,\nu)\leq\zeta+d(T,U)\}\\&\quad\cup\{z\in S:\delta(z,\bar{\nu})\leq\zeta+d(T,U)\}).\end{aligned}\tag{5}$$

On the other hand, set $U=V$ in (2.5)(c). So $\boldsymbol{Q}_U(-1,\alpha+d(T,U))$ and $\boldsymbol{Q}_T(-1,\alpha)$, $\boldsymbol{Q}_U(\nu,\zeta+d(T,U))$ and $\boldsymbol{Q}_T(\nu,\zeta)$, $\boldsymbol{Q}_U(\bar{\nu},\zeta+d(T,U))$ and $\boldsymbol{Q}_T(\bar{\nu},\zeta)$, $\boldsymbol{Q}_U(-1,\beta+d(T,U))$ and $\boldsymbol{Q}_T(-1,\beta)$, are coupled pairs of projections. (6)

Hence, by (5),

$$\begin{aligned}\boldsymbol{R}_U=\boldsymbol{Q}_U(-1,\alpha+d(T,U))|K_T(-1,\alpha)&\oplus\boldsymbol{Q}_U(\nu,\zeta+d(T,U))|K_T(\nu,\zeta)\\&\oplus\boldsymbol{Q}_U(\bar{\nu},\zeta+d(T,U))|K_T(\bar{\nu},\zeta)\end{aligned}\tag{7}$$

is an isomorphism of $K_T(-1,\beta)$ with $K_U(-1,\beta+d(T,U))$, which restricts to isomorphisms of the summands $K_T(-1,\alpha)$ with $K_U(-1,\alpha+d(T,U))$ and so on. However,

$$\begin{aligned}&K_T(\nu,\zeta)\oplus K_T(\bar{\nu},\zeta)=K_T(-1,\beta)\cap K_T(-1,\alpha)^{\perp}\text{ and}\\&\boldsymbol{Q}_U(\nu,\zeta+d(T,U))J=J\boldsymbol{Q}_U(\bar{\nu},\zeta+d(T,U));\end{aligned}$$

that is, $\boldsymbol{Q}_U(\nu,\zeta+d(T,U))$ and $\boldsymbol{Q}_U(\bar{\nu},\zeta+d(T,U))$ are complex conjugates of each other. Consequently (see (2.3)) $\boldsymbol{Q}_U(\nu,\zeta+d(T,U))\oplus\boldsymbol{Q}_U(\bar{\nu},\zeta+d(T,U))$ transfers the standard orientation on $K_T(-1,\beta)\cap K_T(-1,\alpha)^{\perp}$ to the standard orientation on $K_U(-1,\beta+d(T,U))\cap K_U(-1,\alpha+d(T,U))^{\perp}$. Thus $\boldsymbol{R}_U\omega_2$ and $\boldsymbol{Q}_U(-1,\alpha+d(T,U))\omega_1$ are coherent. (8)

Let $p(t)$, $0\leq t\leq 1$, be a path from T to U in $B(T;3\varepsilon)$. For each t, read $p(t)$ instead of U in (7), (8). By (2.5)(b), $\boldsymbol{Q}_{p(t)}(\nu,\zeta+d(T,p(t)))$, $\boldsymbol{Q}_{p(t)}(\bar{\nu},\zeta+d(T,p(t)))$, $\boldsymbol{Q}_{p(t)}(-1,\alpha+d(T,p(t)))$ are continuous in t, so that $\boldsymbol{R}_{p(t)}$ also is, and $\boldsymbol{Q}_T(-1,\beta)\boldsymbol{R}_{p(t)}$, for $0\leq t\leq 1$, is a continuous family of isomorphisms of $K_T(-1,\beta)$ with itself (see (2.5)(c)). Now $\boldsymbol{R}_{p(0)}=\boldsymbol{Q}_T(-1,\beta)$, by virtue of (5); so

$$\boldsymbol{Q}_T(-1,\beta)\boldsymbol{R}_U\omega_2=\boldsymbol{Q}_T(-1,\beta)\omega_2=\omega_2.\tag{9}$$

Take $V=T$ in (2.7)(a). It follows that

$$\begin{aligned}\boldsymbol{R}_U\omega_2 &= \boldsymbol{Q}_U(-1, \beta+d(T, U))\boldsymbol{Q}_T(-1, \beta)\boldsymbol{R}_U\omega_2\\ &= \boldsymbol{Q}_U(-1, \beta+d(T, U))\omega_2.\end{aligned}$$

In view of (8), this establishes the Proposition.

(2.10) COROLLARY. *Suppose in* (2.9) *that* V, $W\in B(T, \varepsilon)$. *Then* $\boldsymbol{Q}_W(-1, \alpha+d(T, W))$ *and* $\boldsymbol{Q}_W(-1, \beta+d(T, W))$ *transform coherent orientations of* $K_V(-1, \alpha+d(T, V))$ *and* $K_V(-1, \beta+d(T, V))$ *into coherent orientations of* $K_W(-1, \alpha+d(T, W))$ *and* $K_W(-1, \beta+d(T, W))$ *respectively.*

PROOF. Apply (2.9) twice (with V, and then W, in place of U) and (2.6) twice (for β itself and for α).

§ 3. Paths in $\boldsymbol{SOC}(E)$.

Throughout this section $\mathfrak{S}$ denotes $\boldsymbol{SOC}(E)$.

(3.1) LEMMA. *Let* $H:[0, 1]\times[0, 1]\longrightarrow\mathfrak{S}$ *be continuous, and let* $\mathscr{U}$ *be an open cover of* $H([0, 1]\times[0, 1])$ *in* $\mathfrak{S}$. *Then there exist finite sequences of numbers,* $0=t_0<t_1<\cdots<t_n=1$ *and* $0=s_0<s_1<\cdots<s_m=1$, *and of triples* $(T_k, \beta_k, \varepsilon_k)$, $1\le k\le l$, *where, for each* k, $T_k\in\mathfrak{S}$, $\beta_k\ge 0$, $\varepsilon_k>0$, $\beta_k+3\varepsilon_k<\pi$, *and*

$$\begin{aligned}\sigma(T_k)\cap\{z\in S:\delta(-1, z)\le\beta_k\}\\ =\sigma(T_k)\cap\{z\in S:\delta(-1, z)\le\beta_k+6\varepsilon_k\},\end{aligned}\tag{1}$$

such that, for $0\le i<n$, $0\le j<m$, *there exists* k, *with* $1\le k\le l$, *and there exists* $U\in\mathscr{U}$, *for which*

$$H([t_i, t_{i+1}]\times[s_j, s_{j+1}])\subseteq B(T_k\,;\,\varepsilon_k)\subseteq U.\tag{2}$$

PROOF. For any $T\in\mathfrak{S}$, $\sigma(T)$ is closed under complex conjugation and has 1 as its only cluster point. Thus there are many possible choices of $\varepsilon>0$ and $\beta\ge 0$ such that $\beta+3\varepsilon<\pi$ and (2.9)(1) is satisfied—for instance β may be 0 and $6\varepsilon=\inf\{\delta(z, -1): z\in\sigma(T),\ z\neq -1\}$. Let $\varkappa$ be a Lebesgue number for the covering $\mathscr{U}$ of $H([0, 1]\times[0, 1])$, and, for each $T\in\mathfrak{S}$, choose values ε_T and β_T for ε and τ so that $\varepsilon_T\le\varkappa$ and (2.9)(1) is satisfied. Now one has an open cover $\{B(T\,;\,\varepsilon_T): T\in\mathfrak{S}\}$ of $\mathfrak{S}$, and the result follows by compactness.

(3.2) Given $T\in\mathfrak{S}$, write $\mathscr{R}(T)=\ker(T+I)$. This is a real subspace of E of even finite dimension (see (1.2)). Let $p(t)$, $0\le t\le 1$, be a continuous path in $\mathfrak{S}$.

(1) The path p will be described here as admissible if it satisfies the conditions that

(a) $(\forall t \in [0, 1]) \mathscr{R}(p(t)) \neq 0$,
(b) $\mathscr{R}(p(0)) = \mathscr{R}(p(1))$.

(2) By an admissible cover of p, refining an open cover of $p([0, 1])$, I shall mean a finite sequence of quadruples $\{(T_i, \beta_i, \varepsilon_i, t_i) : 1 \leq i \leq n\}$, where $0 = t_0 < t_1 < t_2 < \cdots < t_n = 1$ and, for $1 \leq i \leq n$, $T_i \in \mathfrak{G}$, $\beta_i \geq 0$, $\varepsilon_i > 0$, $\beta_i + 3\varepsilon_i < \pi$, $\sigma(T_i) \cap \{z \in S : \delta(-1, z) \leq \beta_i\} = \sigma(T_i) \cap \{z \in S : \delta(-1, z) \leq \beta_i + 6\varepsilon_i\}$ and $p([t_{i-1}, t_i]) \subseteq B(T_i, \varepsilon_i) \subseteq U$ for some $U \in \mathscr{U}$.

(3.3) LEMMA. *Any admissible path has an admissible cover.*

PROOF. In (3.1), take $H(t, s) = p(t)$ for each s, t.

(3.4) Let p be an admissible path; suppose that $\{(T_i, \beta_i, \varepsilon_i, t_i) : 1 \leq i \leq n\}$ is an admissible cover of p. Take an orientation $\omega = \omega(0)$ of $\mathscr{R}(p(0))$. These data induce an orientation $\omega(t)$ on $\mathscr{R}(p(t))$ for each t, as follows.

Let $\omega(t)$ be defined for $0 \leq t \leq t_i < 1$ (where $i \geq 0$). Then $\omega(t_i)$ orients $\mathscr{R}(p(t_i)) = K_{p(t_i)}(-1, 0)$ (see (2.4)). Let ω_i denote the coherent orientation (see (2.8)) of $K_{p(t_i)}(-1, \beta_i + d(T_i, p(t_i))$. For a given $t \in [t_i, t_{i+1}]$, $\boldsymbol{Q}_{p(t)}(-1, \beta_i + d(T_i, p(t)))\omega_i$ is an orientation on $K_{p(t)}(-1, \beta_i + d(T_i, p(t)))$; define $\omega(t)$ to be the coherent orientation on $\mathscr{R}(p(t)) = K_{p(t)}(-1, 0)$. By (2.5)(c), this does indeed extend the definition of $\omega(t)$ for $t_i \leq t \leq t_{i+1}$. In these circumstances, one has

(3.5) LEMMA. *The orientation $\omega(t)$ on $\mathscr{R}(p(t))$ does not depend on the choice of admissible cover of the admissible path p (but only on the orientation ω of $\mathscr{R}(p(0))$).*

PROOF. Say that an admissible cover $\{(T_i, \beta_i, \varepsilon_i, t_i) : 1 \leq i \leq n\}$ of p is refined by another, $\{(T_j', \beta_j', \varepsilon_j', t_j') : 1 \leq i \leq n'\}$, when, for each j, there is an i such that

$$B(T_j'; \varepsilon_j') \subseteq B(T_i; \varepsilon_i), \quad \varepsilon_j' \leq \varepsilon_i, \quad [t_{j-1}', t_j'] \subseteq [t_{i-1}, t_i]. \tag{1}$$

Suppose in these circumstances that the orientation ω on $\mathscr{R}(p(0))$ induces via $\{(T_i, \beta_i, \varepsilon_i, t_i) : 1 \leq i \leq n\}$ the orientation $\omega(t)$ on $\mathscr{R}(p(t))$, and via the refinement $\{(T_j', \beta_j', \varepsilon_j', t_j') : 1 \leq i \leq n'\}$ induces $\omega'(t)$. Assume $\omega(t) = \omega'(t)$ for $0 \leq t \leq t_{j-1}'$, and $[t_{j-1}', t_j'] \subseteq [t_{i-1}, t_i]$. It follows from (2.9) and (2.6), applied to $B(T_i, \varepsilon_i)$, that $\omega(t) = \omega'(t)$ for $t_{j-1}' \leq t \leq t_j'$ also.

To complete the proof, note that any two admissible covers possess a common admissible refinement, by (3.1).

NOTE. The requirement that $\varepsilon_j' \leq \varepsilon_i$ in (1) is superfluous. Since $3\varepsilon_i < \pi$, $3\varepsilon_j' < \pi$ (by definition), (7.2) of [1] ensures that, if $\varepsilon_j' > \varepsilon_i$, there exists a minimising geodesic segment in $B(T_j'; \varepsilon_j')$ of length greater than $2\varepsilon_i$

but less than π. Such a geodesic segment cannot exist in $B(T_i; \varepsilon_i)$. Observe too that, to use (2.6) here, I had to allow nonzero values of β and prove (2.9).

(3.6) Given an admissible path p in $\mathfrak{G}$, take an orientation ω of $\mathscr{R}(p(0))$, and construct the induced orientation $\omega(1)$ on $\mathscr{R}(p(1))=\mathscr{R}(p(0))$, as in (3.4). Define the twist $\rho(p)$ of p to be $+1$ if $\omega(0)=\omega=\omega(1)$, and to be -1 if $\omega\neq\omega(1)$. The choice of ω is evidently irrelevant. As for the condition (3.2)(1)(a), it is needed for the technical purpose of "remembering" the orientation, and may be circumvented as follows.

(3.7) Suppose $p:[0,1]\longrightarrow\mathfrak{G}$ is a path which satisfies the condition (3.2)(1)(b). Let $\boldsymbol{C}^{2n}$ be given its usual complex conjugation, where $n\geq 1$. Take the Hilbert direct sum $E_n=E\oplus C^{2n}$, with the conjugate-isometry $J_n:(x, y)\longmapsto(Jx, \bar{y})$ for $x\in E$, $y\in\boldsymbol{C}^{2n}$. Thus (E_n, J_n) is a real Hilbert space; set, for $0\leq t\leq 1$,

$$p_n(t)=p(t)\oplus(-I_{2n}),$$

where I_{2n} denotes the identity of $\boldsymbol{C}^{2n}$. Now p_n is evidently a continuous path in $\boldsymbol{OC}(E_n)$, and indeed in $\boldsymbol{SOC}(E_n)$ (see (1.2); this is the reason for taking the extra summand to be even-dimensional), and it satisfies both (3.2)(1)(a) and (3.2)(1)(b). Hence $\rho(p_n)$ is defined, and it is easily verified from the definition that it is independent of $n\geq 1$, and agrees with $\rho(p)$ when p satisfies (3.2)(1)(a). Therefore I define $\rho(p)$ to be the common value of $\rho(p_n)$ for positive n; this is the desired extension of the previous definition.

It is clear that the domain of parameters for p (so far taken to be $[0, 1]$) is irrelevant to these arguments.

(3.8) Lemma. *If $\mathscr{R}(p(t))=\mathscr{R}(p(0))=\mathscr{R}$ for all $t\in[0,1]$, then $\rho(p)=0$.*

Proof. As in (3.7), one may assume $\mathscr{R}\neq 0$. Then take, as is evidently possible, an admissible cover of p such that (in the notation of (3.2)(2)) $T_i=p(t_i)$ and $\beta_i=0$ for each i. The projections $\boldsymbol{Q}_{p(t)}(-1, \beta_i+d(p(t_i), p(t_{i+1}))$ of (3.4) are all the orthogonal projection on $\mathscr{R}$. It follows that $\omega(1)=\omega(0)$.

(3.9) Proposition. *Suppose $a, b\in\mathfrak{G}$, and $\mathscr{R}(a)=\mathscr{R}(b)$. Let p and q be paths from a to b which are homotopic in the class of such paths. Then $\rho(p)=\rho(q)$.*

Proof. Let H be a homotopy, as in the hypothesis. As in (3.7), I may assume $\mathscr{R}(\mathrm{H}(s, t))\neq 0$ for all (s, t), by taking the Hilbert direct sum with $-I_2$ if necessary. In the conclusion of (3.1), put $p_i(t)=H(t, s_i)$ for $0\leq i\leq m$. Then $p_0=p$, $p_m=q$. Take an orientation ω on $\mathscr{R}(a)$; given i, with $0\leq i<m$, and l, with $0\leq l\leq n$, define $r_{il}(t)$, $0\leq t\leq 1+s_{i+1}-s_i$, by the prescription:

$$
\begin{aligned}
&r_{il}(t)=p_i(t) \quad \text{for} \quad 0\leq t\leq t_{n-1},\\
&r_{il}(t)=H(t_{n-1}, s_i+t-t_{n-1}) \quad \text{for} \quad t_{n-1}\leq t\leq t_{n-1}+s_{i+1}-s_i,\\
&r_{il}(t)=p_{i+1}(t-s_{i+1}+s_i) \quad \text{for} \quad t_{n-1}+s_{i+1}-s_i\leq t\leq 1+s_{i+1}-s_i.
\end{aligned}
$$

Now $r_{il}(t)$ and $r_{i,l+1}(t)$ are the same for $t\leq t_{n-l-1}$, so r_{il} and $r_{i,l+1}$ transport ω to the same orientation on $\mathscr{R}(p_i(t_{n-l-1}))$. For $t_{n-l-1}\leq t\leq t_{n-1}+s_{i+1}-s_i$, both r_{il} and $r_{i,l+1}$ take values in the same $B(T_k\,;\varepsilon_k)$ (in the notation of (3.1)), and therefore (2.6) ensures that they both induce the same orientation on $\mathscr{R}(p_{i+1}(t_{n-l}))$. For the remaining values of t, r_{il} and $r_{i,l+1}$ once again agree; thus both transport ω to the same orientation of $\mathscr{R}(b)$, and $\rho(r_{il})=\rho(r_{i,l+1})$. Inductively, then, $\rho(r_{i0})=\rho(r_{in})$. But it is trivial that $\rho(p_i)=\rho(r_{i0})$, $\rho(p_{i+1})=\rho(r_{in})$; consequently $\rho(p_0)=\rho(p_m)$ by induction on i.

(3.10) LEMMA. *Let $p, q:[0, 1]\longrightarrow\mathfrak{G}$ be paths such that $p(1)=q(0)$, $\mathscr{R}(p(0))=\mathscr{R}(p(1))=\mathscr{R}(q(1))$. Then, if $*$ denotes concatenation, $\rho(p*q)=\rho(p)\rho(q)$.*

PROOF. The orientation is transported from $\mathscr{R}(p(0))$ to $\mathscr{R}(p(1))$ and thence to $\mathscr{R}(q(1))$.

(3.11) When $p:[0, 1]\longrightarrow\mathfrak{G}$ is a loop (that is, $p(0)=p(1)$), define its translate p^s, for $0\leq s\leq 1$, by $p^s(t)=p(s+t)$ for $t\leq 1-s$, $p^s(t)=p(s+t-1)$ for $t\geq 1-s$. When $x\in\mathfrak{G}$, define xp by setting $(xp)(t)=x\cdot p(t)$ for all t. In these circumstances one has the following Lemma.

LEMMA. (a) $\rho(p^s)=\rho(p)$ *for each* $s\in[0, 1]$.
(b) $\rho(xp)=\rho(p)$ *for each* $x\in\mathfrak{G}$.

PROOF. (a) Define $q_s(t)=p(t)$ for $0\leq t\leq s$. Then $q_s*p^s=p*q_s$, and the result follows by (3.10).

(b) $\mathfrak{G}$ is pathconnected (see (1.2)). Let $r(t)$, $0\leq t\leq 1$, be a path from $p(0)$ to $x\cdot p(0)$; the reverse path is $\bar{r}$. Then p and $r*xp*\bar{r}$ are homotopic with fixed end-points; so, by (3.10), $\rho(p)=\rho(r)\rho(xp)\rho(\bar{r})=\rho(xp)$.

(3.12) REMARKS. (1) The intuitive significance of the preceding arguments is roughly this. When $p(t)$ is a path in $\mathfrak{G}$, the eigenvalues of $p(t)$ vary continuously in t in the sense that, although they may coalesce or bifurcate, they must do so with conservation of multiplicity. Suppose it were possible, as this might suggest, to construct a complete family of continuously varying eigenvectors for $p(t)$, in complex-conjugate pairs. When the eigenvalues corresponding to such a pair pass through -1, the branch which previously had eigenvalue in the upper half-plane may subsequently also have eigenvalue above the real axis; in this case the passage through -1 may be expected to be homotopically trivial. Alternatively, each eigenvector branch may as it were change from one side to the other of the real axis. This reverses the standard orientation (see (2.8)) and inserts a factor -1 in $\rho(p)$.

This description is inadequate, partly because many eigenvalue branches may coalesce simultaneously, but mainly because one cannot in general construct continuously varying eigenvectors across junctions and forks of eigenvalue branches. The formal arguments avoid these obstacles.

(2) As a specific example, consider $\boldsymbol{SO}(2)$. Here $E=\boldsymbol{C}^2$, with the usual conjugation; $\boldsymbol{SO}(2)$ consists of the matrices

$$T(\theta)=\begin{pmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix}, \text{ for real parameters } \theta.$$

The eigenvalues of $T(\theta)$ are $\exp(i\theta)$, $\exp(-i\theta)$, with corresponding eigenvectors $\begin{pmatrix}1\\ i\end{pmatrix}$, $\begin{pmatrix}1\\ -i\end{pmatrix}$ respectively. It is easily checked that

$$d(T(\theta),\ T(\varphi))=\delta(\exp(i\theta),\ \exp(i\varphi)).$$

Let $p(t)=T(t)$ for $0\leq t\leq 2\pi$. This is a loop at e. As in (3.7), one works with $p_1(t)$, which also has constant eigenvectors; it follows that the spectral projections used in (3.4) (and elsewhere) reduce in this case to identities. For $t<\pi$, $F(0, \pi-t, p_1(t))$—see (2.8)—is the span of $\begin{pmatrix}1\\ i\end{pmatrix}\in E$. For $t>\pi$, $F(0, t-\pi, p_1(t))$ is the span of $\begin{pmatrix}1\\ -i\end{pmatrix}$. The standard orientation on $K_{p(t)}(-1, |t-\pi|)\cap K_{p(t)}(-1, 0)^{\perp}$ therefore reverses as t passes through π, and $\rho(p)=-1$. The computation is unchanged if $SO(2)$ is embedded in $\boldsymbol{SO}(n)$, where $n\geq 3$, or in $\boldsymbol{SOC}(E)$, in the customary way, by taking direct sums with identity operators.

In fact the existence of constant eigenvectors for $\boldsymbol{SO}(2)$ enables one to define the "degree" of any loop at e as the number of times (counted with signs) that the eigenvalue associated with $\begin{pmatrix}1\\ i\end{pmatrix}$ passes through -1 in the anticlockwise direction. This integer completely classifies the loop's homotopy type, and its parity gives the twist. However, this procedure is special to $\boldsymbol{SO}(2)$; compare § 4.

(3) The loop p in $\boldsymbol{SO}(2)$, studied above, represents a generator of $\pi_1(\boldsymbol{SO}(2))=\boldsymbol{Z}$; on embedding in $\boldsymbol{SO}(n)$ for $n\geq 3$, or in $\boldsymbol{SOC}(E)$, it represents a generator of $\pi_1(\boldsymbol{SO}(n))=\boldsymbol{Z}_2$ or of $\pi_1(\boldsymbol{SOC}(E))=\boldsymbol{Z}_2$. Therefore, the twist of any loop which represents a generator of $\pi_1(\boldsymbol{SO}(n))$, for $n\geq 2$, or of $\pi_1(\boldsymbol{SOC}(E))$, must be -1 (see (3.9)).

(3.13) THEOREM. *Let $a, b\in\mathfrak{G}$, and $\mathscr{R}(a)=\mathscr{R}(b)$. Two paths $p, q:[0, 1]\longrightarrow\mathfrak{G}$ from a to b are homotopic with fixed end-points if and only if $\rho(p)=\rho(q)$.*

PROOF. Necessity is given by (3.9). Suppose $\rho(p)=\rho(q)$. Then, if $r=a^{-1}(p*\bar{q})$ (see (3.10) and (3.11); $\bar{q}$ is the reverse of q), r is a loop at e and, by (3.11), (3.10), $\rho(r)=\rho(p)-\rho(q)=0$. Hence r does not represent a generator of $\pi_1(G)$, by (3.12)(3), and must be nullhomotopic. This suffices to prove the result.

§ 4. Paths in $UC(E)$.

For the Fredholm unitary group $\boldsymbol{UC}(E)$, the arguments of §§ 2, 3 have very much simpler analogues, which I shall present only in outline. Take $\mathfrak{G}=\boldsymbol{UC}(E)$. For $T\in\mathfrak{G}$, define $\mathscr{R}(T)$ as at (3.2), and say a path $p:[0, 1]\longrightarrow\mathfrak{G}$ is admissible if $\mathscr{R}(p(0))=\mathscr{R}(p(1))$. Admissible covers are defined as at (3.2)(2).

Given an admissible path p, with admissible cover $\{(T_i, \beta_i, \varepsilon_i, t_i):0\leq i\leq n\}$, define for $1\leq i\leq n$

$$\begin{aligned} a_i=&\operatorname{rank} \boldsymbol{P}_{p(t_i)}(\{z\in S:\operatorname{Im} z<0 \text{ and } \delta(z, -1)\leq\beta_i+d(T_i, p(t_i))\})\\ &-\operatorname{rank} \boldsymbol{P}_{p(t_{i-1})}(\{z\in S:\operatorname{Im} z<0 \text{ and } \delta(z, -1)\leq\beta_i+d(T_i, p(t_{i-1}))\}),\end{aligned}$$

and let the degree $\Delta(p)$ of p, relative to the given admissible cover, be the integer $\sum_{i=1}^{n}a_i$. One now finds, by arguments similar to those of §§ 2, 3, that $\Delta(p)$ does not in fact depend on the choice of admissible cover, and is invariant under homotopies with fixed end-points. (One may describe $\Delta(p)$ in intuitive terms as the total algebraic multiplicity of eigenvalues that have passed downwards through -1.) In $\boldsymbol{U}(1)=\boldsymbol{SO}(2)$, the loop p of (3.12)(2) has $\Delta(p)=1$. So, as in (3.13), two admissible paths in $\mathfrak{G}$ between the same points are homotopic with fixed end-points if and only if they have the same degree. More generally, two admissible paths from I to the same point in $\mathfrak{G}$ represent the same point of its n-fold cover if and only if their degrees are congruent mod n.

An alternative construction of the degree was pointed out to me by E. N.

Dancer. By a small deformation of p, one may ensure that -1 is a simple eigenvalue of $p(t)$ for each t such that $p(t)+I$ is not invertible. Now the elements of $\mathfrak{G}$ for which -1 is a simple eigenvalue form an oriented locally closed C^{ω} submanifold of codimension 1 in $\mathfrak{G}$; so, by a second (smaller) deformation of p, one may ensure its intersections with this submanifold are all transversal, whilst preserving the property that -1 is never an eigenvalue of multiplicity greater than 1. The degree is now the appropriate sum of intersection numbers, which does not depend on the choice of the small deformations used. Unfortunately, all the steps I have outlined require technical justification, albeit of a rather familiar kind; such a procedure seems more difficult for $\boldsymbol{SOC}(E)$, since there -1 cannot be a simple eigenvalue; and the proof of (6.6) appears to require something closer to my definition.

§ 5. The Finsler geometry of *SpinC*(E).

In this section, let $\mathfrak{G}$ denote $\boldsymbol{SOC}(E)$ and $\tilde{\mathfrak{G}}$ stand for $\boldsymbol{SpinC}(E)$; the canonical projection is $\boldsymbol{\pi}: \tilde{\mathfrak{G}} \longrightarrow \mathfrak{G}$. Recall (1.3) and (1.4), and, in particular, the path space $P(\mathfrak{G})$. (My treatment here will be much less formal). As in [1], define, for $T \in \mathfrak{G}$,

$$N(T)=\sup\{|\theta| : \theta \in [-\pi, \pi] \text{ and } \exp(i\theta) \in \sigma(T)\}.$$

(5.1) LEMMA. *If $T, U \in \mathfrak{G}$, then $d(T, U) \geq |N(T)-N(U)|$.*

PROOF. By (6.3) of [1], $N(T)=d(T, I)$, $N(U)=d(U, I)$.

(5.2) LEMMA. *If $\tilde{x}, \tilde{y} \in \tilde{\mathfrak{G}}$, then $d(\tilde{x}, \tilde{y}) \geq d(\boldsymbol{\pi}(\tilde{x}), \boldsymbol{\pi}(\tilde{y}))$.*

PROOF. $\boldsymbol{\pi}$ preserves the length of paths (see (1.2)).

(5.3) Because d is both left- and right-invariant, it is easy to see that, whenever $p(t)$, $q(t)$ (for $0 \leq t \leq 1$) are rectifiable paths in $\mathfrak{G}$, the pointwise composition

$$(pq^{-1})(t)=p(t)q(t)^{-1}, \quad 0 \leq t \leq 1,$$

is also rectifiable, with length not exceeding $\ell(p)+\ell(q)$. Thus $P(\mathfrak{G})$ has a group structure which makes both the terminal evaluation $\tau: P(\mathfrak{G}) \longrightarrow \mathfrak{G}$ and its lifting $\tilde{\tau}\boldsymbol{\pi}_*^{-1}: P(\mathfrak{G}) \longrightarrow \tilde{\mathfrak{G}}$ (see (1.3)) into group epimorphisms. Furthermore, if $p, q, pq \in P(\mathfrak{G})$ all satisfy (3.2)(1)(b), then

$$\rho(pq)=\rho(p)\rho(q). \tag{1}$$

Indeed, pq and $p*[\tau(p)q]$ are homotopic with fixed end-points, so (1) follows from (3.9), (3.10), (3.11).

A path $p \in P(\mathfrak{G})$ may be of three kinds. If $\mathscr{R}(\tau(p)) \neq 0$, no twist is defined for p; whilst if $\mathscr{R}(\tau(p))=0$, the twist may be either $+1$ or -1. These three possibilities respect homotopy, and so define a partition of $\tilde{\mathfrak{G}}$ into three disjoint subsets B, A_1, A_{-1} respectively. Evidently $\tilde{e}\in A_1$, $B^{-1}=B$, $A_1^{-1}=A_1$, $A_{-1}^{-1}=A_{-1}$.

The kernel of $\boldsymbol{\pi}$ (when E is of infinite or of odd finite dimension, this is the whole centre of $\tilde{\mathfrak{G}}$) has two elements $\tilde{e}$ and $\tilde{f}$, where $\tilde{f}$ is represented in $P(\mathfrak{G})$ by any noncontractible loop at the identity I of $\mathfrak{G}$. Then, by (1) and (3.13), $\tilde{f}A_1=A_{-1}=A_1\tilde{f}$ and $\tilde{f}B=B=B\tilde{f}$. Consequently $\boldsymbol{\pi}$ is one-to-one on A_1 and on A_{-1}, but two-to-one on B. Of course

$$\begin{aligned}B&=\boldsymbol{\pi}^{-1}\{T\in\mathfrak{G}: -1\in\sigma(T)\}\\ &=\boldsymbol{\pi}^{-1}\{T\in\mathfrak{G}: N(T)=\boldsymbol{\pi}\}.\end{aligned}$$

(5.4) LEMMA. (a) *A_1 and A_{-1} are open and contractible in $\tilde{\mathfrak{G}}$, and* (*b*) *their common frontier is B. Furthermore,* (*c*) *both A_1 and A_{-1} are mapped diffeomorphically by $\boldsymbol{\pi}$ on to $\{T\in\mathfrak{G}: -1\notin\sigma(T)\}=Q$.*

PROOF. Certainly $A_1\cup A_{-1}=\tilde{G}\backslash B=\boldsymbol{\pi}^{-1}(Q)$. But Q is open and contractible in $\mathfrak{G}$ (one may contract along geodesics; see (5.2) of [1]), so that $\boldsymbol{\pi}^{-1}(Q)$ is the disjoint union of two open contractible subsets $\tilde{Q}_1$, $\tilde{Q}_2$ of $\tilde{\mathfrak{G}}$, each mapped diffeomorphically on Q by $\boldsymbol{\pi}$. Any path in Q has twist $+1$, by (3.8) or (3.9); thus every point of $\tilde{Q}_i$ has the same twist for its representatives in $P(\mathfrak{G})$. It follows that $\tilde{Q}_1$, $\tilde{Q}_2$ must be A_1, A_{-1} in some order, and this gives both (a) and (c).

Suppose $T\in\mathfrak{G}$ and $-1\in\sigma(T)$. Thus $\mathscr{R}(T)$ is of non-zero even dimension (see (1.2)) and has an orthonormal basis $e_1, Je_1, e_2, Je_2, \cdots, e_m, Je_m$. Define, for $t\in\boldsymbol{R}$,

$$\begin{aligned}&q(t)|\mathscr{R}(T)^{\perp}=I|\mathscr{R}(T)^{\perp},\\ &q(t)e_j=\exp(it^2)e_j,\ q(t)Je_j=\exp(-it^2)Je_j \quad \text{for } 2\leq j\leq m,\\ &q(t)e_1=\exp(it)e_1,\ q(t)Je_1=\exp(-it)Je_1.\end{aligned}$$

Now $Tq(t)$, for $-\varepsilon\leq t\leq+\varepsilon$ (where $0<\varepsilon<\sqrt{\pi}$), has $\mathscr{R}(Tq(-\varepsilon))=\mathscr{R}(Tq(\varepsilon))=0$ and twist -1 (as in (3.12)). If $p\in P(\mathfrak{G})$ represents $\tilde{x}\in B$, where $\boldsymbol{\pi}(\tilde{x})=p(1)=T$, then, concatenating p with $Tq(t)$ and with $Tq(-t)$ for $0\leq t\leq\varepsilon$, one obtains paths for which twists are defined and opposite (by (3.9), (3.10)). These paths therefore represent points in $\tilde{\mathfrak{G}}$ of which one belongs to A_1 and one to A_{-1}; by choice of ε, these points may be made arbitrarily close to $\tilde{x}$. This proves (b).

(5.5) THEOREM. *If $\tilde{x}\in A_1\cup B$, then $d(\tilde{x}, \tilde{e})=d(\boldsymbol{\pi}(\tilde{x}), I)=N(\boldsymbol{\pi}(\tilde{x}))$. If $\tilde{x}\in A_{-1}$, then $d(\tilde{x}, \tilde{e})=2\boldsymbol{\pi}-N(\boldsymbol{\pi}(\tilde{x}))$.*

PROOF. Suppose $\tilde{x}\in A_1$. By (5.2), $d(\tilde{x}, \tilde{e})\geq d(\boldsymbol{\pi}(\tilde{x}), I)$, and, by (6.3) of [1], $d(\boldsymbol{\pi}(\tilde{x}), I)=N(\boldsymbol{\pi}(\tilde{x}))$. Conversely, the unique minimising geodesic from I to $\boldsymbol{\pi}(\tilde{x})$ (see (7.2), (7.3) of [1]) has length $N(\boldsymbol{\pi}(\tilde{x}))$ and lifts to a geodesic from $\tilde{e}$ to $\tilde{x}$, by virtue of (5.4)(c); thus, as required, $d(\tilde{x}, \tilde{e})\leq N(\boldsymbol{\pi}(\tilde{x}))$ also.

If $\tilde{x}\in B$, $d(\tilde{x}, \tilde{e})\geq N(\boldsymbol{\pi}(\tilde{x}))=\pi$. But, by (5.4)(b), $\tilde{x}$ may be approximated by elements of A_1, whose distance from $\tilde{e}$ does not exceed π. Thus $d(\tilde{x}, \tilde{e})=\pi$.

Now suppose $\tilde{x}\in A_{-1}$. Any rectifiable path from $\tilde{x}$ to $\tilde{e}$ must pass through B, by (5.4)(b). Let $\tilde{b}$ be a point in B on the path. The length of the path cannot be less than

$$\begin{aligned} d(\tilde{x}, \tilde{b})+d(\tilde{b}, \tilde{e}) &\geq |N(\pi(\tilde{b}))-N(\pi(\tilde{x}))|+\pi \\ &=2\pi-N(\pi(\tilde{x})). \end{aligned}$$

To establish the opposite inequality, let me suppose $\pi(\tilde{x})=T$, where $N(T)<\pi$, and let η be a unit eigenvector for the eigenvalue $\exp(iN(T))$ of T. Let

$$S=\log(T|\{\eta, J\eta\}^{\perp}) \quad \text{(see (5.1), (5.2) of [1])},$$

and define $p(t)$, for $0\leq t\leq 1$, by

$$\begin{aligned} &p(t)|\{\eta, J\eta\}^{\perp}=\exp(tS), \\ &p(t)J\eta=\exp(i(2\pi-N(T))t\}J\eta, \\ &p(t)\eta=\exp\{-i(2\pi-N(T))t\}\eta. \end{aligned}$$

Thus p joins I to T in $\mathfrak{G}$, and has twist -1 (compare (3.12)). Hence it represents $\tilde{x}$. However, $\ell(p)=2\pi-N(T)$ trivially. It follows (see (1.4)) that $d(\tilde{x}, \tilde{e})\leq 2\pi-N(T)$, as required.

(5.6) COROLLARY. *For any $\tilde{x}, \tilde{y}\in\tilde{\mathfrak{G}}$, $d(\tilde{x}, \tilde{y})+d(\tilde{x}, \tilde{y}\tilde{f})=2\pi$. In particular, $d(\tilde{x}, \tilde{y})=2\pi$ if and only if $\tilde{x}=\tilde{y}\tilde{f}$.*

(5.7) For the statement and proof of the next theorem, it is helpful to introduce a special notation. Given $\tilde{x}, \tilde{y}\in\tilde{\mathfrak{G}}$, let $F(\tilde{x}, \tilde{y})$ denote the sum of the eigenspaces in E of the operator $\boldsymbol{\pi}(\tilde{y}^{-1}\tilde{x})$ which correspond to the eigenvalues $\exp(\pm iN(\boldsymbol{\pi}(\tilde{y}^{-1}\tilde{x})))$.

(5.8) THEOREM. (a) *Any two points $\tilde{x}, \tilde{y}\in\tilde{\mathfrak{G}}$ may be joined by a minimising geodesic.*

(b) *If $p(t)$, $q(t)$ (for $0\leq t\leq 1$) are minimising geodesics joining $\tilde{x}$ to $\tilde{y}$, then*

$$(\forall t\in[0, 1]) \ \boldsymbol{\pi}(p(t))|F(\tilde{x}, \tilde{y})^{\perp}=\boldsymbol{\pi}(q(t))|F(\tilde{x}, \tilde{y})^{\perp}.$$

(*Compare* (7.3) *of* [1].)

(c) *There is only one minimising geodesic between* $\tilde{x}$ *and* $\tilde{y}$ *when* $\tilde{y}^{-1}\tilde{x}\in A_1$, *or when* $\tilde{y}^{-1}\tilde{x}\in B$ *and* $\dim F(\tilde{x}, \tilde{y})=2$. *When* $\tilde{y}^{-1}\tilde{x}\in A_{-1}$ *and* $\dim F(\tilde{x}, \tilde{y})=2$, *there are exactly two minimising geodesics between* $\tilde{x}$ *and* $\tilde{y}$. *In all other cases there are infinitely many.*

PROOF. It suffices to assume $\tilde{y}=\tilde{e}$. In the proof of (5.5), I constructed minimising geodesics between $\tilde{e}$ and $\tilde{x}$ when $\tilde{x}\in A_1$ and when $\tilde{x}\in A_{-1}$. Suppose now that $\tilde{x}\in B$, and let $\boldsymbol{\pi}(\tilde{x})=T$. Take an orthonormal basis $e_1, Je_1, \cdots, e_m, Je_m$ for $\mathscr{R}(T)=\mathscr{R}$, and define, for $0\leq t\leq 1$,

$$r(t)|\mathscr{R}^{\perp}=\exp(t\log(T|\mathscr{R}^{\perp})) \quad \text{(see (5.1) of [1])},$$

$$r(t)e_j=\exp(i\pi t)e_j,\ r(t)Je_j=\exp(-i\pi t)Je_j$$

for each j, $1\leq j\leq m$; and similarly

$$s(t)|\{e_1, Je_1\}^{\perp}=r(t)|\{e_1, Je_1\}^{\perp},$$
$$s(t)e_1=\exp(-i\pi t)e_1,\ s(t)Je_1=\exp(i\pi t)Je_1.$$

Then $r(t)$ and $s(t)$ are both minimising geodesics from I to T in $\mathfrak{G}$ (compare (5.4), (6.2) of [1]). But $r*\bar{s}$ is a loop at I whose twist is -1 (as in (3.12)); thus, from (3.9), r and s must represent distinct points of $\tilde{\mathfrak{G}}$, and one of them must represent $\tilde{x}$. The lifting of this path to $\tilde{\mathfrak{G}}$ is the required minimising geodesic from $\tilde{e}$ to $\tilde{x}$. This completes the proof of (a).

Suppose $\tilde{x}\in A_1$, and $p(t)$ is a minimising geodesic from $\tilde{e}$ to $\tilde{x}$. Then, by virtue of (5.5), $\boldsymbol{\pi}(p(t))$ is a minimising geodesic from I to $\boldsymbol{\pi}(\tilde{x})=T$ in $\mathfrak{G}$. Since $N(T)<\pi$, $\mathscr{R}(T)=0$, and $\boldsymbol{\pi}(p(t))$ is uniquely determined, by (7.3) of [1]. If $\tilde{x}\in B$, and $p(t), q(t)$ are minimising geodesics from $\tilde{e}$ to $\tilde{x}$, then $\boldsymbol{\pi}(p(t))$ and $\boldsymbol{\pi}(q(t))$ are in the same way minimising geodesics in $\mathfrak{G}$; and $\boldsymbol{\pi}(p(t))|F(\tilde{x}, \tilde{e})^{\perp}=\pi(q(t))|F(\tilde{x}, \tilde{e})^{\perp}$ by (7.3) of [1].

Thirdly, suppose that $\tilde{x}\in A_{-1}$ and $\boldsymbol{\pi}(\tilde{x})=T$ as before. If $p(t)$, $0\leq t\leq 1$, is a minimising geodesic from $\tilde{e}$ to $\tilde{x}$ in $\tilde{\mathfrak{G}}$, then there exists a compact real skew-adjoint operator S in E such that $\boldsymbol{\pi}p(t)=\exp(tS)$ for $0\leq t\leq 1$ and $\|S\|=2\pi-N(T)$. Certainly S and T commute, so that $F(\tilde{x}, \tilde{e})^{\perp}=H$ is S-invariant, $\exp(S|H)=T|H$. Set $\lambda=\|S|H\|$; then $i\lambda$ is an eigenvalue of $S|H$ and $\exp(i\lambda)$ is an eigenvalue of $T|H$. Consequently either $\lambda\leq N(T|H)$ or $\lambda\geq 2\pi-N(T|H)$. But $N(T|H)<N(T)$, and $\lambda=\|S|H\|\leq\|S\|=2\pi-N(T)$, so that in fact $\|S|H\|\leq N(T|H)$. It follows that $\exp(t(S|H))$, for $0\leq t\leq 1$, is a minimising geodesic in $\boldsymbol{SOC}(H)$ between $I|H$ and $T|H$, and $S|H=\log(T|H)$ (see (5.2) and (7.3) of [1]). This completes the proof of (b). Now suppose in addition that $F(\tilde{x}, \tilde{e})$ is of dimension 2. Then $S|F(\tilde{x}, \tilde{e})$ must be real, skew-adjoint, with eigenvalues $\pm i(2\pi-$

$N(T))$ (since, as I have shown, its eigenvalues on H have smaller absolute values), and there are exactly two such operators in $F(\tilde{x}, \tilde{e})$. If the dimension of $F(\tilde{x}, \tilde{e})$ (which must of course be even) is greater than 2, there is an infinite family of real skew-adjoint operators in $F(\tilde{x}, \tilde{e})$ for which the eigenvalues $\pm i(2\pi - N(T))$ have equal odd multiplicity and all other eigenvalues are $\pm iN(T)$. Each such operator is a candidate for $S|F(\tilde{x}, \tilde{e})$, so that, as stated, there are in this case infinitely many minimising geodesics from $\tilde{e}$ to $\tilde{x}$ in $\tilde{\mathfrak{G}}$.

If $\tilde{x} \in B$ and $F(\tilde{x}, \tilde{e})$ is of dimension 2, the same analysis as for the previous case shows that any minimising geodesic from $\tilde{e}$ to $\tilde{x}$ must project on to one of the paths r, s constructed previously; indeed, if the geodesic is $p(t)$, then $p(t) = \exp(tS)$, where S is fixed on $F(\tilde{x}, \tilde{e})^{\perp} = \mathscr{R}(T)^{\perp}$ (as before, $\boldsymbol{\pi}(\tilde{x}) = T$) and $S|F(\tilde{x}, \tilde{e})$ is real skew-adjoint with eigenvalues $\pm i\pi$. However, r and s represent different points of B, as above. So there is only one minimising geodesic between $\tilde{e}$ and $\tilde{x}$ in $\tilde{\mathfrak{G}}$. On the other hand, if $F(\tilde{x}, \tilde{e})$ is of dimension greater than 2, there are infinitely many bases in it which may be used to construct paths r and s; therefore there are infinitely many minimising geodesics.

(5.9) LEMMA. *Between any two points in $\tilde{\mathfrak{G}}$ there are infinitely many minimising paths which are not geodesics.*

PROOF. It suffices to assume the points are close together, and then to apply the analogous result in $\mathfrak{G}$ (see (7.4) and (7.8) of [1]).

§ 6. The universal cover of $UC(E)$.

Henceforth, let $\mathfrak{G}$ denote $\boldsymbol{UC}(E)$; $\tilde{\mathfrak{G}}_n$ is its n-fold covering group and $\tilde{\mathfrak{G}}$ its universal covering group. The projections are $\boldsymbol{\pi} : \tilde{\mathfrak{G}} \longrightarrow \mathfrak{G}$ and $\boldsymbol{\pi}_n : \tilde{\mathfrak{G}}_n \longrightarrow \mathfrak{G}$, the identities are $\tilde{e}$ and $\tilde{e}_n$. In this context, (5.1) remains true, and so do the analogues of (5.2) for $\tilde{\mathfrak{G}}_n$ and for $\tilde{\mathfrak{G}}$.

As before, $P(\mathfrak{G})$ has a group structure, and, if p, q, pq are admissible elements of $P(\mathfrak{G})$ in the sense of § 4, then

$$\Delta(pq) = \Delta(p) + \Delta(q). \qquad \cdots\cdots\cdots\cdots (*)$$

(Compare (5.3) and § 4.) I shall discuss $\tilde{\mathfrak{G}}$ first.

(6.1) A path $p \in P(\mathfrak{G})$ for which $\mathscr{R}(\tau(p)) \neq 0$ has no degree. If $\mathscr{R}(\tau(p)) = 0$, the degree may be any integer. Thus $\tilde{\mathfrak{G}}$ is partitioned into disjoint subsets: for any $j \in \boldsymbol{Z}$, A_j consists of those points represented by paths of $P(\mathfrak{G})$ which have degree j, and B consists of the points whose representative paths are not admissible—that is

$$B=\boldsymbol{\pi}^{-1}\{T\in\mathfrak{G}:\mathscr{R}(T)\neq 0\}.$$

The kernel of $\boldsymbol{\pi}$ is isomorphic to $\boldsymbol{Z}$, and the degree of representative loops furnishes a specific isomorphism. Let $\tilde{f}_j$ be that element of the kernel represented by loops of degree j. Evidently $(*)$ implies that $\tilde{f}_k A_j=A_{j+k}$ and that $\tilde{f}_k B=B$ for all $j, k\in\boldsymbol{Z}$. The elements $\tilde{f}_k$ are central in $\tilde{\mathfrak{G}}$; when E is infinite-dimensional, they constitute the whole centre of $\tilde{\mathfrak{G}}$, since the centre of $\mathfrak{G}$ is trivial.

(6.2) THEOREM. (a) *For each $j\in\boldsymbol{Z}$, A_j is open and contractible as a subset of $\tilde{\mathfrak{G}}$, and is mapped diffeomorphically by $\boldsymbol{\pi}$ on to $\{T\in\mathfrak{G}:-1\notin\sigma(T)\}$.*

(b) *$\{A_j:j\in\boldsymbol{Z}\}$ is a locally finite family of sets in $\tilde{\mathfrak{G}}$.*

(c) *Let $\tilde{x}\in B$. Then $\{i:\tilde{x}\in\bar{A}_i\}$ is an interval $[\ell,\ell+k]$ in $\boldsymbol{Z}$, where $k=\dim\mathscr{R}(\boldsymbol{\pi}(\tilde{x}))$. (Note that this remains true for any element of $\tilde{\mathfrak{G}}$.)*

PROOF. (a) Copy the proof of (5.4)(a), (c).

For (b) and (c), let $\tilde{x}\in B$ and $T=\boldsymbol{\pi}(\tilde{x})$. Thus there is an $\varepsilon\in(0,\pi)$ such that $\{z\in S:\delta(z,-1)<6\varepsilon\}\cap\sigma(T)=\{-1\}$. Hence, if $\tilde{y}\in B$ and $d(\tilde{x},\tilde{y})<\varepsilon$, $\boldsymbol{\pi}(\tilde{y})=U$, then, by (5.2) above, $d(T,U)<\varepsilon$, and (4.7) of [1] (see also (2.5) above) gives

$$\operatorname{rank}\boldsymbol{P}_U(\{z\in S:\delta(z,-1)\leq d(T,U)\})=\operatorname{rank}\boldsymbol{P}_T(\{-1\})=k.$$

Consequently, if $\tilde{y},\tilde{z}\in B(\tilde{x};\varepsilon)\backslash B$ and p is a path from $\tilde{y}$ to $\tilde{z}$ lying in $B(\tilde{x};\varepsilon)$ (such paths certainly exist, by the definition of d), then $|\Delta(p)|\leq k$. In view of $(*)$, this means that, if $\tilde{y}\in A_j$ and $\tilde{z}\in A_i$, $|j-i|\leq k$, which establishes (b).

(A) Let $e_1, e_2,\cdots,e_k$ be an orthonormal basis for $\mathscr{R}(T)$. Given n, with $0\leq n\leq k$, and $0\leq t\leq\gamma$, define

$$q_n(t)e_j=\exp(it)e_j \quad\text{for}\quad 1\leq j\leq n,$$
$$q_n(t)e_j=\exp(-it)e_j \quad\text{for}\quad n<j\leq k, \text{ and}$$
$$q_n(t)|\mathscr{R}(T)^{\perp}=I|\mathscr{R}(T)^{\perp}.$$

On compounding Tq_m, Tq_n respectively with a path $p\in P(\mathfrak{G})$ representing $\tilde{x}$, one obtains paths p_m, p_n representing points $\tilde{y}_m$, $\tilde{y}_n$ of $B(\tilde{x};\gamma)$. The paths p_m, p_n are clearly admissible in the sense of § 4, and therefore $\tilde{y}_m\in A_i$ and $\tilde{y}_n\in A_j$ for some $i,j\in\boldsymbol{Z}$. Now

$$\begin{aligned}\Delta(p_m)-\Delta(p_n)&=\Delta(p*Tq_m)-\Delta(p*Tq_n)\\&=\Delta(T\bar{q}_n*\bar{p}*p*Tq_m)\\&=\Delta(T\bar{q}_n*Tq_m)\quad\text{(homotopy invariance)}\\&=m-n,\end{aligned}$$

from the definitions of Δ in § 4 and of q_m, q_n above. Consequently

$$i-j=m-n\,;$$

and, since γ may be arbitrarily small, this proves that x lies in the closure of $A_\ell, A_{\ell+1}, \cdots, A_{\ell+k}$, where ℓ is defined by the statement that $\tilde{y}_0 \in A_\ell$. Since it was proved above that $B(x\,;\,\varepsilon)$ meets at most $k+1$ of the sets A_h, this completes the demonstration of (c).

(6.3) In view of (6.2), $\tilde{\mathfrak{G}}$ is partitioned into subsets $B_{r,s}$, for $r, s \in \boldsymbol{Z}$ and $s \geq r$, where $\tilde{x} \in B_{r,s}$ if and only if it is in the closure of A_r, $A_{r+1}, \cdots$, A_s and of no other A_h. The interval $[r, s]$ is in some sense a generalised degree of $\tilde{x}$, with indeterminacy $s-r=\dim \mathscr{R}(\boldsymbol{\pi}(\tilde{x}))$. Now $B_{r,r}=A_r$ and is open. In general, it follows from (6.2) that, for given $m, n \in \boldsymbol{Z}$ with $n \geq m$, the set

$$\bigcup_{[r,s]\subseteq[m,n]} B_{r,s}$$

is open.

(Whilst this partition of $\tilde{\mathfrak{G}}$ is analogous to that given for $\boldsymbol{SpinC}(E)$ at (5.3), it is less natural. For $\boldsymbol{SpinC}(E)$, the eigenvalue -1 has a certain canonical status as the only real element of $S\backslash\{1\}$. For $\boldsymbol{UC}(E)$, one might use any $\lambda \in S\backslash\{1\}$ instead of -1, both in § 4 and in (6.2). This would alter the class of admissible paths, and the new degree would in general differ from the old when both were defined - they are both defined, and agree, for loops at I. The corresponding change in the sets $B_{r,s}$ would, however, be a hindrance in what follows.)

(6.4) Given $\theta \in [-2\pi, 2\pi]$, define

$$I(\theta)=\{\exp[i(\pi+t\theta)] : 0<t\leq 1\}.$$

Then, for $0 \neq m \in \boldsymbol{Z}$ and $T \in \boldsymbol{UC}(E)$, set

$$N_m(T)=\inf\{x>0 : \text{rank } \boldsymbol{P}_T(I(mx/|m|)) \geq |m|\}.$$

(That is, for $m>0$ one considers spectral projections corresponding to segments of S of length x extending anticlockwise from -1; for $m<0$, the segments are to extend clockwise.) When E is of infinite dimension, rank $\boldsymbol{P}_T(I(\theta))$ is infinite for $\theta>\pi$, so that $N_m(T) \leq \pi$. Furthermore, $N_m(T)$ must be an attained infimum if it is strictly less than π; if $N_m(T)=\pi$, it may or may not be attained.

(6.5) NOTE. $B_{r,s}^{-1}=B_{-s,-r}$ and $N_m(T)=N_{-m}(T^{-1})$.

(6.6) PROPOSITION. *Let $p(t)$, $0 \leq t < a$, be a rectifiable path in $\tilde{\mathfrak{G}}$ which* *starts at $\tilde{e}$ and is parametrised by arc-length. Suppose v is an odd positive integer, and $p(v\pi)\in B_{k,\ell}$. Then, for any $t\in[v\pi, \min(a, (v+2)\pi))$, if $p(t)\in B_{r,s}$,*

$$\operatorname{rank} \boldsymbol{P}_{\boldsymbol{\pi}p(t)}(\{\exp(i\theta): v\pi<\theta\leq t\})\geq r-k. \tag{1}$$

In particular, if $t\in[\pi, \min(a, 3\pi))$,

$$\operatorname{rank} \boldsymbol{P}_{\boldsymbol{\pi}p(t)}(\{\exp(i\theta): \pi<\theta\leq t\})\geq r. \tag{2}$$

PROOF. First I prove (2) when $t=\pi$. By (5.2), $d(\boldsymbol{\pi}p(t), I)\leq d(p(t), \tilde{e})$; when $0\leq t<\pi$, then, (6.1) of [1] shows that $-1\notin\sigma(\boldsymbol{\pi}p(t))$, and so, by (6.2)(a) above, the lifting $p(t)$ of $\boldsymbol{\pi}p(t)$ lies in A_0. By (6.3), then, $p(\pi)\in B_{r,s}$ where $r\leq 0\leq s$, and (2) holds for $t=\pi$ (and for $0\leq t<\pi$). It will now be sufficient to prove (1).

Suppose that τ_0 is the infimum of the values of $t\in[v\pi, \min(a, (v+2)\pi))$ for which (1) is untrue. Choose $\varepsilon>0$ such that

$$\sigma(\boldsymbol{\pi}p(\tau_0))\cap\{z\in S: \delta(z, -1)\leq 28\varepsilon\}\subseteq\{-1\} \tag{3}$$

and

$$\min(a, (v+2)\pi)-\tau_0\geq 5\varepsilon. \tag{4}$$

Take $\tau\in[\max(v\pi, \tau_0-\varepsilon), \tau_0]$ such that (1) holds for $t=\tau$. (If possible, one may take $\tau=\tau_0$). In turn, suppose henceforth that $\tau\leq t\leq\tau+2\varepsilon$. Thus

$$d(p(t), p(\tau_0))\leq 2\varepsilon,\ d(p(t), p(\tau))\leq 2\varepsilon,\ d(p(\tau), p(\tau_0))\leq\varepsilon. \tag{5}$$

At (6.2)(A), one sees how to approximate $p(\tau)\in B_{i,j}$ and $p(t)\in B_{m,n}$ along paths of length less than ε by $\tilde{x}\in A_i$ and $\tilde{y}\in A_m$ respectively, in such a way that (see (6.4))

$$\boldsymbol{P}_{\boldsymbol{\pi}p(\tau)}(I(2\varepsilon))=\boldsymbol{P}_{\boldsymbol{\pi}(\tilde{x})}(I(2\varepsilon)) \tag{6}$$

and

$$\boldsymbol{P}_{\boldsymbol{\pi}p(\tau)}(I(2\varepsilon+d(p(t), p(\tau))))=\boldsymbol{P}_{\boldsymbol{\pi}(\tilde{y})}(I(2\varepsilon+d(p(t), p(\tau))))\,; \tag{7}$$

for the construction (6.2)(A) moves the eigenvalue -1 a little clockwise without changing the rest of the spectrum. From (5),

$$d(\tilde{x}, p(\tau_0))<2\varepsilon,\ d(\tilde{x}, \tilde{y})<4\varepsilon. \tag{8}$$

Now apply (2.5)(7), and (5.2), to (3) and (8). One finds

$$\sigma(\boldsymbol{\pi}(\tilde{x}))\cap\{z\in S: \delta(z, -1)\leq 2\varepsilon\}=\sigma(\boldsymbol{\pi}(\tilde{x}))\cap\{z\in S: \delta(z, -1)\leq 26\varepsilon\}. \tag{9}$$

Since $26\varepsilon-2\varepsilon>6d(\boldsymbol{\pi}(\tilde{x}), \boldsymbol{\pi}(\tilde{y}))$, apply § 4 (as in the proof of (6.2)(c)), and obtain, in view of (6) and (7),

$$m-i=\operatorname{rank} \boldsymbol{P}_{\pi p(t)}(I(2\varepsilon+d(p(t), p(\tau))))-\operatorname{rank} \boldsymbol{P}_{\pi p(\tau)}(I(2\varepsilon)). \tag{10}$$

Again by (2.5)(7), with (3) and (5),

$$\sigma(\boldsymbol{\pi}p(t))\cap\{z\in S: \delta(z,-1)\leq 2\varepsilon\}=\sigma(\boldsymbol{\pi}p(t))\cap\{z\in S: \delta(z,-1)\leq 26\varepsilon\}, \tag{11}$$

$$\sigma(\boldsymbol{\pi}p(\tau))\cap\{z\in S: \delta(z,-1)\leq \varepsilon\}=\sigma(\boldsymbol{\pi}p(\tau))\cap\{z\in S: \delta(z,-1)\leq 27\varepsilon\}. \tag{12}$$

From (11) and (12) respectively, one deduces

$$\boldsymbol{P}_{\pi p(t)}(I(2\varepsilon+d(p(t), p(\tau))))=\boldsymbol{P}_{\pi p(t)}(I(2\varepsilon)), \tag{13}$$

$$\boldsymbol{P}_{\pi p(\tau)}(I(2\varepsilon))=\boldsymbol{P}_{\pi p(\tau)}(I(\varepsilon)). \tag{14}$$

Suppose $t\geq v\pi+26\varepsilon$, so $\tau\geq t-2\varepsilon>v\pi+24\varepsilon$; and set $\alpha=(\tau-v\pi-24\varepsilon)/2$ and $\mu=\exp[i(v\pi+24\varepsilon+\tau)/2]$. Thus

$$\{z\in S: \delta(z,\mu)\leq\alpha\}=\{\exp(i\theta): v\pi+24\varepsilon\leq\theta\leq\tau\}$$

and

$$\{z\in S: \delta(z,\mu)\leq\alpha+t-\tau\}=\{\exp(i\theta): v\pi+24\varepsilon-t+\tau\leq\theta\leq t\}.$$

The length of $\boldsymbol{\pi}p$ between the parameter values τ and t does not exceed $t-\tau$. Hence one may use (4.9) of [1]:

$$\begin{aligned}&\operatorname{rank} \boldsymbol{P}_{\pi p(\tau)}(\{\exp(i\theta): v\pi+24\varepsilon\leq\theta\leq\tau\})\\ &\quad\leq\operatorname{rank} \boldsymbol{P}_{\pi p(t)}(\{\exp(i\theta): v\pi+24\varepsilon-t+\tau\leq\theta\leq t\}).\end{aligned} \tag{15}$$

However, $v\pi+24\varepsilon-t+\tau\geq v\pi+22\varepsilon$ and, from (4), $t+3\varepsilon\leq\min(a,(v+2)\pi)$. Consequently (11) and (12) yield

$$\begin{aligned}&\boldsymbol{P}_{\pi p(\tau)}(\{\exp(i\theta): v\pi<\theta\leq\tau\})\\ &\quad=\boldsymbol{P}_{\pi p(\tau)}(I(\varepsilon))+\boldsymbol{P}_{\pi p(\tau)}(\{\exp(i\theta): v\pi+24\varepsilon\leq\theta\leq\tau\}),\end{aligned} \tag{16}$$

$$\begin{aligned}&\boldsymbol{P}_{\pi p(t)}(\{\exp(i\theta): v\pi<\theta\leq t\})\\ &\quad=\boldsymbol{P}_{\pi p(t)}(I(2\varepsilon))+\boldsymbol{P}_{\pi p(t)}(\{\exp(i\theta): v\pi+22\varepsilon\leq\theta\leq t\}),\end{aligned} \tag{17}$$

where the sums are direct. Between the second terms on the right, (15) shows the rank does not decrease; (10), (13) and (14) prove that the first terms increase in rank by $m-i$ exactly. Since (1) holds for τ, it follows that it holds for t also. The same conclusion is valid if $t\leq v\pi+26\varepsilon$; in that case the second terms on the right of (16) and (17) do not appear, in view again of (11) and (12), and (10), (13), and (14) deal with the first terms as before. Now t was any point of $[\tau,\tau+2\varepsilon]\supseteq[\tau_0,\tau_0+\varepsilon]$, so the definition of τ_0 has led to a contradiction, and the Proposition is proved.

(6.7) THEOREM. *Suppose E is of infinite dimension, and $\tilde{x}\in B_{k,\ell}\subseteq\tilde{\mathfrak{G}}$. Then*

(a) *if* $k\leq 0\leq\ell$, $d(\tilde{x},\tilde{e})=N(\boldsymbol{\pi}(\tilde{x}))$ (*see* §5, *or* (3.3) *of* [1]);

(b) *if* $0<k$, $d(\tilde{x}, \tilde{e})=\pi+N_k(\boldsymbol{\pi}(\tilde{x}))$ (*see* (6.4));
(c) *if* $\ell<0$, $d(\tilde{x}, \tilde{e})=\pi+N_\ell(\boldsymbol{\pi}(\tilde{x}))$.

PROOF. Suppose first that $k=0=\ell$. Then $\tilde{x}\in A_0$. Each path in A_0 is the lifting of a path of the same length in $\{T\in\mathfrak{G}: -1\notin\sigma(T)\}$, and a path from $\tilde{x}$ to $\tilde{e}$ which leaves A_0 must similarly be of length at least π (see (6.3) of [1]). So $d(\tilde{x}, \tilde{e})=d(\boldsymbol{\pi}(\tilde{x}), I)=N(\boldsymbol{\pi}(\tilde{x}))$, as required.

If $k\leq 0\leq \ell\neq k$, then, again by (6.3) of [1], $d(\tilde{x}, \tilde{e})\geq d(\boldsymbol{\pi}(\tilde{x}), I)=N(\boldsymbol{\pi}(\tilde{x}))=\pi$; however, by the definition (6.3), $\tilde{x}\in\bar{A}_0$ and so $d(\tilde{x}, \tilde{e})\leq\pi$ also. This proves (a). Next I prove (b), and (c) will follow by inversion (as in (6.5)).

Suppose, now, that $\tilde{x}\in B_{k,\ell}$ where $0<k$. Take any $\theta\in(N_k(\pi(\tilde{x})), 2\pi)$; then rank $\boldsymbol{P}_{\pi(\tilde{x})}(I(\theta))\geq k$, and one may select k orthonormal eigenvectors $e_1, e_2, \cdots, e_k$ of $\boldsymbol{\pi}(\tilde{x})$, corresponding to eigenvalues $\exp[i(\pi+\nu_1)], \cdots, \exp[i(\pi+\nu_k)]$, where $0<\nu_j\leq\nu_{j+1}<\theta$ for $1\leq j<k$. Let F be the span of $e_1, \cdots, e_k$, and let $\nu\in(0, \pi)$ be such that $\exp[i(\pi+\tau)]$ is not an eigenvalue of $\boldsymbol{\pi}(\tilde{x})$ for $\tau\in(0, \nu)$. Now define a path $p(t)$, $0\leq t\leq 1$, in $\boldsymbol{UC}(E)$:

$$p(t)|F^{\perp}=\exp[(1-t)\log(\boldsymbol{\pi}(\tilde{x})|F^{\perp})] \quad \text{and} \tag{1}$$
$$p(t)e_j=\boldsymbol{\pi}(\tilde{x})\cdot\exp[-it(\pi+\nu_j)]\cdot e_j \quad \text{for} \quad 1\leq j\leq k.$$

(See (5.1) of [1] for the meaning of $\log(\pi(\tilde{x})|F^{\perp})$.) The length of p is $\pi+\nu_k$, since $\|\log(\pi(\tilde{x}))|F^{\perp})\|<\pi$.

Let $\tilde{p}$ denote the lifting of p to $\tilde{\mathfrak{G}}$ which starts at $\tilde{x}$. As at (6.2)(A), one sees that, when $0<\tau<\nu/(\pi+\nu)$, $-1\notin\sigma(p(\tau))$ and $\tilde{p}(\tau)\in A_k$ (the definition of $p(t)$ is such that the eigenvalue -1 of $\boldsymbol{\pi}(\tilde{x})$ is moved clockwise). So $p|[\tau, 1]$ has a degree, which is easily seen to be $-k$; indeed, only $p(t)|F$ contributes to it. It follows that $\tilde{p}(1)\in A_0$, and therefore that $\tilde{p}(1)=\tilde{e}$, since $p(1)=I$. One deduces that

$$d(\tilde{x}, \tilde{e})\leq\ell(\tilde{p})=\ell(p)=\pi+\nu_k<\pi+\theta,$$

and, in view of the arbitrariness of θ, this shows

$$d(\tilde{x}, \tilde{e})\leq\pi+N_k(\boldsymbol{\pi}(\tilde{x})).$$

The converse inequality is an immediate consequence of (6.6).

NOTE. To reconcile the various formulae, observe that a direct proof of (c) would require taking the logarithm of -1 to be $-i\pi$ in (5.1) of [1] and (1) above.

(6.8) THEOREM. *Let E be infinite-dimensional. Suppose $\tilde{x}, \tilde{y}\in\tilde{\mathfrak{G}}$ and $\tilde{z}=\tilde{x}^{-1}\tilde{y}\in B_{k,\ell}$. Then*:

(a) *$\tilde{x}$ and $\tilde{y}$ may be joined by infinitely many geodesics.* (*In particular, $\tilde{\mathfrak{G}}$ is exponential*).

(b) *$\tilde{x}$ and $\tilde{y}$ may be joined by a minimising path if and only if either $k\leq 0\leq \ell$; or $0<k\leq \operatorname{rank} \boldsymbol{P}_{\pi(\tilde{z})}(I(\pi))$; or $0<-\ell\leq \operatorname{rank} \boldsymbol{P}_{\pi(\tilde{z})}(I(-\pi))$. Such a minimising path is unique if and only if $\tilde{x}=\tilde{y}$.*

(c) *When $\tilde{x}$ and $\tilde{y}$ may be joined by a minimising path, they may also be joined by a minimising geodesic.*

(d) *A minimising geodesic between $\tilde{x}$ and $\tilde{y}$ is unique if and only if either 0 is an end-point of $[k, \ell]$; or, if $0<k$, and $\operatorname{rank}\boldsymbol{P}_{\pi(\tilde{z})}(I(N_k(\boldsymbol{\pi}(\tilde{z}))))=k$; or, if $\ell<0$, and $\operatorname{rank} \boldsymbol{P}_{\pi(\tilde{z})}(I(-N_\ell(\pi(\tilde{z}))))=|\ell|$.*

PROOF. It will suffice to suppose $\ell\geq 0$, since the case $\ell<0$ may be treated by inversion. Let $\boldsymbol{\pi}(\tilde{z})=T\in \boldsymbol{UC}(E)$, and let m be a nonnegative integer. Suppose first that $k\geq 0$, and choose orthonormal eigenvectors $e_1, \cdots, e_{k+2m}$ for T, with corresponding eigenvalues $\exp(i\lambda_1), \cdots, \exp(i\lambda_{k+2m})$ such that $\pi<\lambda_j<3\pi$ for each j. This may be done in infinitely many ways, even if $k=0$ (by exploiting the choice of m). Let F be the span of $e_1, \cdots, e_{k+2m}$, and define a path r in $\boldsymbol{UC}(E)$ by setting, for $0\leq t\leq 1$,

$$
\begin{aligned}
r(t)|F^{\perp}&=\exp(t\log(T|F^{\perp})),\\
r(t)e_j&=\exp(it\lambda_j)e_j \quad \text{for} \quad 1<j\leq k \quad \text{and for}\\
&\qquad\qquad j=k+1, k+3, \cdots, k+2m-1,\\
r(t)e_{k+2j}&=\exp(it(\lambda_{k+2j}-4\pi))e_{k+2j} \text{ for } 1\leq j\leq m.
\end{aligned}
$$

Exactly as for the path p constructed in (6.7), r must lift to a path from $\tilde{e}$ to $\tilde{z}$. Indeed, $r(t)|F^{\perp}$ passes no eigenvalues through -1, and approaches -1 anticlockwise; while $r(t)|F$ has the appropriate degree k, the eigenvectors e_{k+2j}, e_{k+2j-1} contributing -1, $+1$ respectively for $1\leq j\leq m$.

Now suppose $k\leq 0\leq \ell$; choose orthonormal eigenvectors $e_1, \cdots, e_{\ell-k}$, $e_{\ell-k+1}, \cdots, e_{\ell-k+2m}$ for T, such that $e_1, \cdots, e_{\ell-k}$ span $\ell(T)$ and the eigenvalue corresponding to $e_{\ell-k+j}$ is $\mu_j\in(-\pi, \pi)$, for $1\leq j\leq 2m$. Again this choice may be made in infinitely many ways, even if $k=0=\ell$. Define

$$
\begin{aligned}
s(t)|F^{\perp}&=\exp(t\log(T|F^{\perp})),\\
s(t)e_j&=\exp(i\pi t)e_j \quad \text{for} \quad 1\leq j\leq \ell, \qquad\qquad (1)\\
s(t)e_j&=\exp(-i\pi t)e_j \quad \text{for} \quad \ell<j\leq \ell-k,\\
s(t)e_{\ell-k+2j-1}&=\exp(it(2\pi+\mu_{2j-1}))e_{\ell-k+2j-1} \quad \text{for} \quad 1\leq j\leq m,\\
s(t)e_{\ell-k+2j}&=\exp(it(\mu_{2j}-2\pi))e_{\ell-k+2j} \quad \text{for} \quad 1\leq j\leq m.
\end{aligned}
$$

If $\tilde{s}$ is the lifting of s which ends at $\tilde{z}$, then $\tilde{s}(t)\in A_0$ for t close to 1. This is clear, since, if (1) held for $1\leq j\leq \ell-k$ instead, one would have

$\tilde{s}(t)\in A_k$ for t close to 1 (compare the proof of (6.2)(c)). Hence $\tilde{s}$ joins $\tilde{e}$ to $\tilde{z}$. This completes the proof of (a), as both $\tilde{r}$ and $\tilde{s}$ are geodesics.

Since $d(\tilde{z}, \tilde{e})\leq\pi$ by (6.7), (6.6) shows that the condition stated in (b) is necessary. If $k\leq 0\leq \ell$, take $m=0$ in the definition of $\tilde{s}$ above to obtain a minimising geodesic from $\tilde{e}$ to $\tilde{z}$. If $k>0$, the condition asserts that $N_k(T)$ is an attained infimum (see (6.4)); if it is satisfied, one may therefore choose $e_1, \cdots, e_k$ in the definition of r so that $\pi<\lambda_j\leq N_k(T)$ for $1\leq j\leq k$, and set $m=0$. Then $\tilde{r}$ is a minimising geodesic from $\tilde{e}$ to $\tilde{z}$, as required. When $\tilde{x}\neq\tilde{y}$, minimising paths between them cannot be unique; even short segments may be altered in infinitely many ways, since the same is true in $\mathfrak{G}$ by (7.8) of [1]. Thus (b) and (c) are both established.

Suppose that $k<0<\ell$. The above construction of the minimising geodesic $\tilde{s}$ from $\tilde{e}$ to $\tilde{z}$ (with $m=0$) may then be modified by other choices of $e_1, \cdots, e_{\ell-k}$. Similarly, if $0<k$ and rank $\boldsymbol{P}_T(I(N_k(T)))>k$, the choice of $e_1, \cdots, e_k$ in the definition of $\tilde{r}$ is non-unique. In both cases there are infinitely many possible choices. So the conditions of (d) are necessary. Suppose they are satisfied.

Let $q(t)=\exp(tS)$, $0\leq t\leq 1$, be the projection on $\mathfrak{G}$ of a minimising geodesic $\tilde{q}$ from $\tilde{e}$ to $\tilde{z}$ in $\tilde{\mathfrak{G}}$. Here S is compact skew-adjoint (so is completely determined by its eigenvalues and eigenvectors), $\|S\|=d(\tilde{e}, \tilde{z})$, and $\exp S=T$. Each eigenspace $F_\theta=\ker(T-\exp(i\theta))$, for $0\leq\theta<2\pi$, is invariant under S, and the eigenvalues of $S|F_\theta$ must be logarithms of $\exp(i\theta)$. Take first the case $k>0$. Since $\|S\|=\pi+N_k(T)$ by (6.7), the only admissible eigenvalues for $S|F_\theta$ are:

$$i\theta \text{ when } 0\leq\theta\leq\pi+N_k(T), \tag{1}$$
$$i(\theta-2\pi) \text{ when } \pi-N_k(T)\leq\theta<2\pi, \tag{2}$$

and, in addition, $2\pi i$ when $\theta=0$ and $N_k(T)=\pi$.

Should $S|F_\theta$ have an eigenvalue $-i\pi$ with multiplicity v, then, for τ sufficiently near to 1, $\tilde{q}(\tau)\in A_{k+v}$, and the degree of $q|[0, \tau]$ must be $k+v$. Each eigenvalue of S belonging to $[-i\pi, i\pi]$ contributes zero to the degree; an eigenvalue in $(i\pi, i(\pi+N_k(T))]$ adds in its multiplicity, and one in $[-i(\pi+N_k(T)), -i\pi)$ subtracts off its multiplicity. By hypothesis, the total multiplicity of all the eigenvalues of T which have logarithms in $[i\pi, i(\pi+N_k(T))]$ is exactly k. So $v=0$ necessarily, and the eigenvalues of S must satisfy (1) for all $\theta\in[0, \pi+N_k(T)]$ (including 0 and π), and (2) when $\pi+N_k(T)<\theta$. This fixes S.

When $\ell=0$, $d(\tilde{e}, \tilde{z})=\|S\|\leq\pi$ and the possible eigenvalues of $S|F_\theta$ are: $i\theta$ if $\theta\in[0, \pi)$, $-i(2\pi-\theta)$ if $\theta\in(\pi, 2\pi)$, and, for $\theta=\pi$, $i\pi$ and $-i\pi$. Certainly $\tilde{q}(t)\in A_0$ for all $t\in[0, 1)$, as it must be at distance less than π

from $\tilde{e}$; this is only possible if $S|F_\pi = i\pi I|F_\pi$ (once more, see the proof of (6.2)(c)). Similarly, if $k=0$, $S|F_\pi = -i\pi I|F_\pi$. So again S is uniquely determined.

(6.9) REMARKS. I have shown that for infinite-dimensional E, $\tilde{\mathfrak{G}}$ is exponential and bounded, with diameter exactly 2π ; in fact $d(\tilde{f}_j, \tilde{f}_k) = 2\pi$ if $j \neq k$ (see (6.1)). If E is of countable (Hilbert) dimension, there exist pairs of points in $\tilde{\mathfrak{G}}$ which cannot be joined by minimising paths. Suppose that $\tilde{z} \in B_{k,\prime}$ and $n = \operatorname{rank} \boldsymbol{P}_{\pi(\tilde{z})}(I(\pi)) < \infty$. This condition means that 1 is in the essential spectrum of $\pi(\tilde{z})$ only because it is a cluster point of eigenvalues with positive imaginary part. Then $\tilde{f}_m \tilde{z}$ cannot be joined to $\tilde{e}$ by a minimising path when $m+k>n$. If $m+k \leq n$, there is a minimising path - for $m+k<0$, this results from the observation that $\operatorname{rank} \boldsymbol{P}_{\pi(\tilde{z})}(I(-\pi))$ must be infinite. Since all points for which there is no minimising path to $\tilde{e}$ are obtained either in this way or by subsequent inversion, it is easy to see that they form a set of first category in $\tilde{\mathfrak{G}}$. On the other hand, when E is non-separable, 1 is an eigenvalue of uncountable multiplicity of every operator in $\mathfrak{G}$, and every point of $\tilde{\mathfrak{G}}$ satisfies the conditions of (6.8)(c).

§ 7. Finite coverings and finite dimensions.

(7.1) The n-fold covering $\tilde{\mathfrak{G}}_n$ of $\mathfrak{G} = \boldsymbol{UC}(E)$ may be treated as a quotient of the universal cover $\tilde{\mathfrak{G}}$. The degree mod n of representative loops furnishes a specific isomorphism of $\ker \boldsymbol{\pi}_n$ with $\boldsymbol{Z}_n = \boldsymbol{Z}/(n\boldsymbol{Z})$, under which the projection f_k of $\tilde{f}_k$ (see (6.1)) corresponds to the residue class $<k>$ of k mod n. Describe a subset D of $\boldsymbol{Z}_n$ as an interval if, for any $b, c \in D$, there exists a finite sequence $d_1, \cdots, d_j$ of elements of D, such that $d_1 = b$, $d_j = c$, and, for $1 \leq i < j$, $d_{i+1} = d_i \pm <1>$. Thus singletons and $\boldsymbol{Z}_n$ itself are intervals. Let $|D|$ denote the number of elements of D. One now has, after (6.1) and (6.2):

(7.2) THEOREM. *$\tilde{\mathfrak{G}}_n$ is partitioned into disjoint subsets B_D, one for each non-empty interval D in $\boldsymbol{Z}_n$. If $D = \{d\}$, $B_D = A_d$ consists of those points whose representative paths in $\mathfrak{G}$ are admissible (in the sense of § 4) and have degrees whose residue classes* mod *n are equal to d. Then A_d is open in $\tilde{\mathfrak{G}}_n$, and $\boldsymbol{\pi}_n$ maps A_d diffeomorphically on to $\{T \in \mathfrak{G} : -1 \notin \sigma(T)\}$. When D is not a singleton, B_D consists of those points which lie in the closure of A_d if and only if $d \in D$; and $\bigcup_{C \subseteq D} B_C$ is open (where C ranges over intervals). For each $x \in B_D$,*

$$|D| - 1 = \min(\dim \mathscr{R}(\boldsymbol{\pi}_n(x)), n-1).$$

(7.3) THEOREM. *Let E be of infinite dimension; let $x, y \in \mathfrak{G}_n$, and suppose that $x^{-1}y \in B_D$. Write $T = \boldsymbol{\pi}_n(x^{-1}y)$. Then:*

(a) *if $<0> \in D$, $d(x, y) = N(T)$, whilst, if $<0> \notin D$,*

$$d(x, y) = \pi + \min\{N_k(T) : k \in \boldsymbol{Z} \text{ and } <k> \in D\}; \tag{1}$$

(b) *x and y may be joined by infinitely many geodesics;*

(c) *x and y may be joined by a minimising path if and only if either $<0> \in D$, or, for some $k \in \boldsymbol{Z}$ realising the minimum of $N_k(T)$ for $<k> \in D$ (see (1)),*

$$|k| \leq \text{rank } \boldsymbol{P}_T(I(k\pi/|k|));$$

(d) *if x and y may be joined by a minimising path, they may be joined by a minimising geodesic;*

(e) *a minimising geodesic between x and y is unique if and only if either $<0>$ is an end-point of D, or if*

$$0 < k = \text{rank } \boldsymbol{P}_T(I(N_k(T))), \tag{2}$$

where k is the least integer realising the minimum in (1), or if

$$0 < -k = \text{rank } \boldsymbol{P}_T(I(-N_k(T))), \tag{3}$$

where k is the greatest integer realising the minimum. If the minimum is attained both for a negative and for a positive value of k, where the positive value satisfies (2) and the negative value satisfies (3), there are exactly two minimising geodesics between x and y. In all other cases there are infinitely many.

NOTE. Since $N_k(T)$ increases for positive increasing k, and also for negative decreasing k, the minimum in (1) is really over only two values of k, which are determined by D. Which of them (if not both) gives the minimum depends on T. At most one positive integer can simultaneously realise the minimum in (1) and satisfy (2), and similarly for (1) and (3).

Theorems (7.2), (7.3) follow from (6.2), (6.7), (6.8).

(7.4) (a) When E is real and of finite dimension greater than 2, the arguments of §§ 2, 3, and 5 need no modification; the two-dimensional case is adequately discussed at (3.12)(2).

(b) When E is complex and of positive finite dimension u, § 4 applies without alteration. In § 6, however, infinite-dimensionality is often assumed (in particular from (6.7) onwards); and, in (6.4), $N_m(T)$ is formally infinite when $|m| > u$, is always an attained infimum for $|m| \leq u$, and may take any values in $(0, 2\pi]$.

(7.5) THEOREM. *Let E be complex of finite positive dimension u; let $\tilde{x}$, $\tilde{y}\in\tilde{\mathfrak{G}}$, where $\tilde{z}=\tilde{x}^{-1}\tilde{y}\in B_{r,s}$, and set $T=\boldsymbol{\pi}(\tilde{z})$.*

(a) *If $r\leq 0\leq s$, then $d(\tilde{x}, \tilde{y})=N(T)$.*

(b) *If $0<r=ua+b$, where $a, b\in\boldsymbol{Z}$ and $1\leq b\leq u$, then*

$$d(\tilde{x}, \tilde{y})=(2a+1)\pi+N_b(T).$$

(c) *If $0<-s=ua'+b'$, where $a', b'\in\boldsymbol{Z}$ and $1\leq b'\leq u$, then*

$$d(\tilde{x}, \tilde{y})=(2a'+1)\pi+N_{-b'}(T).$$

PROOF. Compare (6.5), (6.7); I need only prove (b).

Let $e_1, \cdots, e_u$ be a complete orthonormal system of eigenvectors of T, with corresponding eigenvalues $\exp[i(\pi+\nu_1)], \cdots, \exp[i(\pi+\nu_m)]$, ordered so that $0<\nu_1\leq\nu_2\leq\cdots\leq\nu_b=N_b(T)\leq\nu_{b+1}\cdots\leq\nu_m\leq 2\pi$. Define a skew-adjoint operator S in E as follows:

$$Se_j=i[(2a+1)\pi+\nu_j]e_j \quad \text{for} \quad 1\leq j\leq b,$$
$$Se_j=i[(2a-1)\pi+\nu_j]e_j \quad \text{for} \quad b<j\leq m.$$

The path $p(t)=\exp(tS)$ in $\tilde{\mathfrak{G}}$, for $0\leq t\leq 1$, has length $(2a+1)\pi+N_b(T)$. For all τ sufficiently close to 1, though perhaps not for $\tau=1$, $-1\notin\sigma(\exp(\tau S))$ and $p|[0, \tau]$ has degree $au+b$. Indeed, the eigenvalue of e_j passes $a+1$ times through -1 when $1\leq j\leq b$, and a times when $j>b$. So, as in (6.7), p represents $\tilde{z}$ and

$$d(\tilde{x}, \tilde{y})=d(\tilde{e}, \tilde{z})\leq(2a+1)\pi+N_b(T). \tag{1}$$

Given $h\in\boldsymbol{Z}$, let C_h denote the union of the sets $B_{g,f}$ for $h\geq g\in\boldsymbol{Z}$. This set is closed, by (6.3). Take q to be any rectifiable path in $\tilde{\mathfrak{G}}$ which starts at $\tilde{e}$ and is parametrised by arc-length; since such a path may be indefinitely extended, one may assume $q(t)$ is defined for all $t\geq 0$. For $0\leq t<\pi$, certainly $q(t)\in A_0$, and therefore $q(\pi)\in C_0$. Suppose v is an odd positive integer, and $q(v\pi)\in C_{(v-1)u/2}$. Therefore $q(v\pi)\in B_{k,\ell}$, where $k\leq(v-1)u/2$. If $v\pi\leq t<(v+2)\pi$ and $q(t)\in B_{m,j}$, then

$$m-k\leq\operatorname{rank}\boldsymbol{P}_{\pi q(t)}(\{\exp(i\theta): v\pi<\theta\leq t\}), \tag{2}$$

from (6.6). Hence, in particular, $m\leq k+u\leq(v+1)u/2$, and $q(t)\in C_{(v+1)u/2}$, which is closed. So $q((v+1)\pi)\in C_{(v+1)u/2}$ also, and, by induction, $q(t)\in C_{(v-1)u/2}$ for any odd positive integer v and any $t\in[0, v\pi]$. (3)

If $q(t_0)=\tilde{z}$, where, by hypothesis, $\tilde{z}\notin C_{au}$, then (3) shows that $t_0>(2a+1)\pi$. If $t_0\leq(2a+3)\pi$, then state (2), with $v=2a+1$, $i=r$, $t=t_0$, in the form

$$N_b(T)\leq N_{r-k}(T)\leq t_0-(2a+1)\pi,$$

so that $t_0 \geq (2a+1)\pi + N_b(T)$. (Recall (7.4)(b) for the value of $N_b(T)$.) The same holds when $t_0 > (2a+3)\pi$. This establishes the converse inequality to (1).

(7.6) THEOREM. *Retain the hypotheses of* (7.5). *Then*:

(a) $\tilde{x}$ *and* $\tilde{y}$ *may be joined by a minimising geodesic.*

(b) *There is only one minimising geodesic between* $\tilde{x}$ *and* $\tilde{y}$ *if and only if either* 0 *is an end-point of* $[r, s]$; *or, if* $0<r$, *and* $\operatorname{rank} \boldsymbol{P}_T(I(N_b(T)))=b$; *or, if* $0<-s$, *and* $\operatorname{rank} \boldsymbol{P}_T(I(-N_{b'}(T)))=b'$.

(c) *If* $u=1$, *all minimising paths are geodesics. If* $u>1$, *every minimising path between* $\tilde{x}$ *and* $\tilde{y}$ *is a geodesic if and only if either* $r \leq 0 \leq s$ *and all elements of* $\sigma(T)$ *have the same real part, or* $\tilde{z}$ *is central.*

PROOF. The first paragraph of the proof of (7.5) demonstrates (a); and (b) only requires the obvious modifications in the proof of (6.8)(d). Suppose $r \leq 0 \leq s$. Then, by (7.5)(a), any minimising path q from $\tilde{e}$ to $\tilde{z}$ in $\tilde{\mathfrak{G}}$ projects to a minimising path $\boldsymbol{\pi}q$ in $\mathfrak{G}$. By (7.6) of [1], $\boldsymbol{\pi}q$ is necessarily a geodesic (and therefore so is q) if and only if all elements of $\sigma(T)$ have the same real part.

Now suppose $u>1$, and $0<r$. Let $\tilde{z}$ be such that any minimising path from $\tilde{e}$ to $\tilde{z}$ is a geodesic. Take a minimising geodesic $p(t)$ from $\tilde{e}$ to $\tilde{z}$ in $\tilde{\mathfrak{G}}$, and let $2t_0$ be the last value of the parameter such that $p(2t_0) \in \bar{A}_0$. All elements of $\sigma(\boldsymbol{\pi}p(t_0))$ have the same real part, since otherwise $p|[0, t_0]$ may be substituted by a minimising path which is not a geodesic (lifted from $\mathfrak{G}$; again, see (7.6) of [1]), to give a minimising path from $\tilde{e}$ to $\tilde{z}$ which is not geodesic. As $p(2t_0)=(p(t_0))^2$, all elements of $\sigma(\boldsymbol{\pi}p(2t_0))$ have equal real parts, and so, in view of (7.5)(a), $\sigma(\boldsymbol{\pi}p(2t_0))=\{-1\}$ and $p(2t_0)=-I$. If $p|[0, 2t_0]$ does not take only central values, conjugation in $\mathfrak{G}$ gives many other minimising paths between I and $-I$ that are homotopic to it with fixed end-points; lifting to $\tilde{\mathfrak{G}}$ and substituting for $p|[0, 2t_0]$, one obtains non-geodesic minimising paths between $\tilde{e}$ and $\tilde{z}$. Thus $\boldsymbol{\pi}p|[0, 2t_0]$, and consequently $p|[0, 2t_0]$, take only central values in $\mathfrak{G}$ and $\tilde{\mathfrak{G}}$, and $\tilde{z}$ must itself becentral, $T=\exp[i(\pi+\theta)]\cdot I$, where $\theta \in (0, 2\pi]$. Here $N_b(T)=\theta$, and $s=r$, if $\theta \neq 2\pi$, or $s=r+u$, if $\theta=2\pi$. Let $e_1, \cdots, e_u$ be an orthonormal basis in E.

Suppose first that $u \nmid r$. Define a path in $\mathfrak{G}$ as follows.

$$q(t)e_j=\exp(it)e_j \quad \text{for} \quad 0 \leq t \leq (2a+1)\pi+\theta \quad \text{and} \quad 1 \leq j \leq b,$$
$$\text{and for} \quad 0 \leq t \leq 2a\pi \quad \text{and} \quad j>b;$$

whilst, if $j>b$ and $2a \leq t \leq (2a+1)\pi+\theta$,

$$q(t)e_j=\exp[i(\theta-\pi)(t-2a\pi)/(\theta+\pi)]e_j.$$

Then q is a path in $\mathfrak{G}$ of length $d(\tilde{e}, \tilde{z})$; it is uniformly parametrised, begins at I, ends at T, and (if admissible) has degree $r=au+b$. It therefore lifts to a minimising path from $\tilde{e}$ to $\tilde{z}$ in $\tilde{\mathfrak{G}}$, which is not a geodesic. (When q is not admissible—that is, when $\theta=2\pi$—argue as in (7.5)(b).)

As usual, the case when $0<-s$ may be settled by inversion. To conclude, therefore, suppose that $0<r$, $\tilde{z}$ is central, and $u|r$, so that $b=u$. Let $p_1(t)$, $0\leq t\leq\eta=d(\tilde{e}, \tilde{z})$, be the unique minimising geodesic from $\tilde{e}$ to $\tilde{z}$ (see (b)); for each t, $\boldsymbol{\pi}p_1(t)=\exp(it)\cdot I$. Suppose $p(t)$, $0\leq t\leq\eta$, is a minimising path from $\tilde{e}$ to $\tilde{z}$ in $\tilde{\mathfrak{G}}$, and define

$$t_1=\inf\{t\in[0, \eta]: p_1|[t, \eta]=p|[t, \eta]\}.$$

Evidently t_1 is an attained infimum. As the degrees of restrictions of P_1 are sums in which each summand is either 0 or u, certainly $p_1(t_1)\in B_{i,j}$, where $u|i$ and $u|j$. If $t_1\leq\pi$, then $p_1(t_1)\in B_{0,0}\cup B_{0,u}$, and, by the previous cases, $p|[0, t_1]$ is necessarily a geodesic, and in turn (see (b)) necessarily equal to $p_1|[0, t_1]$. So, if $t_1\neq 0$, $t_1>\pi$.

Suppose $t_1\equiv\pi+\theta_1 \pmod{2\pi}$, where $0<\theta_1\leq 2\pi$, and let ε be (for instance) $\theta_1/7$. Take $t\in[t_1-\varepsilon, t_1]$. Then

$$\sigma(\boldsymbol{\pi}p(t_1))=\sigma(\boldsymbol{\pi}p_1(t_1))=\{\exp(it_1)\}.$$

The length of $\boldsymbol{\pi}p$ between parameters t and t_1 is t_1-t, so that, by (4.7) of [1],

$$\sigma(\boldsymbol{\pi}p(t))\subseteq\{\zeta\in S: \delta(\zeta, \exp(it_1))\leq t_1-t\}. \tag{1}$$

Hence, because of the choice of ε, no eigenvalues of $\boldsymbol{\pi}p(t)$ can have passed through -1 in the negative (clockwise) sense as compared with $\boldsymbol{\pi}p(t_1)$; whilst, if any had moved downwards through or from -1, that would have increased i and (by (7.5)(b)) $d(\tilde{e}, p(t))$ could not have decreased. Thus in fact $p(t)\in B_{i,k}$ still (for some k), and by (7.5)(b) again

$$t_1-t=d(\tilde{e}, p(t_1))-d(\tilde{e}, p(t))=N_u(\boldsymbol{\pi}p(t_1))-N_u(\boldsymbol{\pi}p(t)).$$

So

$$\sigma(\boldsymbol{\pi}p(t))\subseteq\{\exp[i(\tau+\pi)]: 0\leq\tau\leq\theta_1-t_1+t\}. \tag{2}$$

Put (1) and (2) together: $\sigma(\boldsymbol{\pi}p(t))=\{\exp(it)\}$, and thus $\pi p(t)=\pi p_1(t)$, for any $t\in[t_1-\varepsilon, t_1]$. This contradicts the definition of t_1; so the hypothesis that $t_1>0$ must be false, and p coincides with p_1 as required.

(7.7) It is easily seen that $\tilde{x}$ and $\tilde{y}$ (in (7.6)) may be joined by infinitely many geodesics. One may also deduce from (7.5) and (7.6) the analogous results for $\tilde{\mathfrak{G}}_n$ (compare (7.2), (7.3)).

References

[1] C. J. ATKIN, The Finsler geometry of groups of isometries of Hilbert space, J. Aust. Math. Soc. **42** (1987), 196-222.

[2] K. D. ELWORTHY, Fredholm maps and $GL_c(E)$-structures, Bull. Amer. Math. Soc. **74** (1968), 582-586.

[3] P. de la HARPE, The Clifford algebra and the spinor group of a Hilbert space, Comp. Math. **25** (1972), 245-261.

[4] P. J. HILTON and S. WYLIE, 'Homology theory', Cambridge University Press, 1960.

[5] R. J. PLYMEN and R. F. STREATER, A model of the universal covering group of $SO(H)_2$, Bull. Lond. Math. Soc. **7** (1975), 283-288.

Department of Mathematics
Victoria University of Wellington