



HAL
open science

Einstein and the old quantum theory

Alexander Afriat

► **To cite this version:**

| Alexander Afriat. Einstein and the old quantum theory. 2019. hal-02424959v2

HAL Id: hal-02424959

<https://hal.science/hal-02424959v2>

Preprint submitted on 14 Jan 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Einstein and the old quantum theory

Alexander Afriat

January 14, 2020

Abstract

Objecting that Sommerfeld's quantum conditions refer to particular coordinates, Einstein proposes a more invariant rule. Even if the invariance is in fact *canonical*, as Graffi (2005) has pointed out, Einstein may have in mind a double *configuration* space invariance: with respect to loop deformations, and to point transformations—all on a torus where features of the Liouville-Arnol'd theorem already appear.

Contents

1	Introduction	1
2	The two quantum rules	3
3	Loops, <i>trajectoires</i>, gauge	5
4	Hamilton-Jacobi-Einstein	7
5	Liouville-Einstein-Arnol'd?	12
6	<i>Einwertigkeit</i>, separability, quantisation	16
7	Final remarks	19

1 Introduction

Einstein's role in quantum theory is well known: one associates him with the foundational debate in the twenties and thirties, with the photoelectric effect, perhaps with the quantum theory of gases—or even statistical mechanics—in general; but less with the old quantum theory of Bohr and Sommerfeld (from 1913), or with analytical mechanics for that matter. “Zum Quantensatz von Sommerfeld und Epstein” (1917a), which I propose to consider, has an unusual place in the history of science, characterised by neglect, limited attention then unexpected, retarded recognition. Gutzwiller (1990) and Graffi (2005) mention the peculiar history of its citations: practically half a century of silence,¹ then rediscovery² in the more mathematical context of analytical mechanics, dynamical systems. Even if analytical mechanics lay outside of Einstein's main mechanical interests—statistical mechanics, relativistic mechanics, foundations of mechanics—*Quantensatz* should be treated more as analytical mechanics than as quantum theory. Sommerfeld's quantum rule³ ((1) below), referred to in the title, becomes little more than a point of departure, almost a pretext for the development or at least adumbration of a futuristic, topological, highly invariant analytical mechanics; that at any rate is what it amounts to, however Einstein himself saw it.

In a nutshell, Einstein (1917a,b) objects that Sommerfeld's rule refers to particular coordinates, in the sense that there's a condition for each of the l coordinates q_1, \dots, q_l . Einstein proposes another rule ((4) below), which is more invariant inasmuch as it integrates the *entire* momentum one-form \mathbf{p} —not one component p_i at a time—over the l homotopy classes that characterise the topology of the torus he eventually introduces (§2 below). Which raises the issue of why Einstein's rule should be any better than Sommerfeld's. *Empirical* superiority is often important in physics. Not here: Einstein never even brings it up, being only concerned with formal invariance; and even for us, favoured as we are by hindsight, empirical adequacy can hardly be invoked to discriminate theories we now know to be empirically very inadequate.⁴ But perhaps there are cases where Einstein's rule makes more sense than Sommerfeld's, or works better. For an elliptical Kepler motion, Einstein's rule is no better than Sommerfeld's, both work well: the score there is *one all*. And neither rule can really handle the self-intersections of a more complicated *Rosettenbahn*: nil nil (cumulatively one all). The superiority of Einstein's rule will only emerge on the toroidal configuration space he constructs to resolve the self-intersections: one nil there (cumulatively two one). But even on the torus, the advantage is only qualified: all one can say is that Einstein's rule makes more

¹Punctuated by two important wave-mechanical citations: Broglie (1924) pp. 63ff, and the footnote on p. 495 of Schrödinger (1926). The story of how the *mécanique ondulatoire* emerged—see Lanczos (1952) pp. 277ff—as if by mistake from Louis de Broglie's remarkable misunderstandings of *Quantensatz* deserves to be told, but elsewhere; and Schrödinger suggests, in the footnote, that *Quantensatz* was behind his *Wellenmechanik* as well.

²See Keller (1958), Gutzwiller (1990) pp. 215, 282, 315, Spivak (2010) p. 640.

³Sommerfeld (1915) p. 429, Sommerfeld (1916) p. 9; see also Epstein (1916a) p. 173, Epstein (1916b) pp. 490, 501.

⁴There is no reason why the two rules should agree on motions, ruling out the same ones. But even if that would make the rules significantly inequivalent, it would only give the preference to one if an independent (empirical) way of checking the selections were available. We now know that an *undulatory* quantum theory—with superposition, coherence, interference, resonance *etc.*—is needed to make any sense at all of atomic mechanics, and select the right motions.

sense and is more invariant, not that Sommerfeld's rule would make no sense at all.⁵

The peculiarities of *Quantensatz*, which is paradoxically both primitive⁶ and decades ahead of its time, give rise to peculiar historiographical difficulties which require a peculiar, *ad hoc* methodology. Some of the mathematics is so futuristic that anachronisms are needed to make any sense of it at all; but there one's on a slippery slope that can lead to the worst exaggerations. Anachronisms, however indispensable in such cases, have to be kept under control, it is too easy to get carried away.

Much of my analysis will concern three central (and indeed related) themes: (i) *mechanics in configuration space* Q (as opposed to phase space Γ); (ii) *integrability, ergodicity*; (iii) *Liouville-Arnol'd*:

- (i) Even if Einstein uses the Hamiltonian function of position *and momentum* (and confusingly makes a reference to *Phasenraum*⁷), his mechanics seems to be firmly rooted in the q_i -*Raum* he so often mentions.⁸ It is not a symplectic mechanics, in phase space.⁹
- (ii) Einstein's treatment of integrability and ergodicity is undeniably modern, perhaps decades ahead of its time. But does it all make sense? Is mere configuration space enough? Would phase space be needed after all?
- (iii) One can easily identify features of the Liouville-Arnol'd theorem (§5 below) in *Quantensatz*. But is it *all* there? Does Einstein really have a *steady* motion on a *rigid* torus? Or just a shapeless, merely 'topological' torus?

Is it possible that Einstein, with little or no symplectic geometry, anticipated so much modern, geometrical mechanics: integrability, ergodicity, Liouville-Arnol'd?

"Integrability" can mean various things. There is the purely geometrical notion, well exemplified by the exactness of a one-form $\alpha = df$: infinitesimal objects $\alpha(\mathbf{q})$ assigned to every point \mathbf{q} fit together in such a way as to allow 'derivation [df] from a potential f ,' as Einstein would put it. But that's abstract; other notions are more concretely arithmetical, and have to do with the simplicity or even the very possibility of coordinate representation. The numbers involved¹⁰ acquire geometrical meaning by identifying manifolds, which can simplify the representation of (integrated) motion by *adaptation*—by 'following it so as to eliminate it,' absorbing its twists & turns into their adapted shapes, thus enclosing and therefore representing it.

Einstein has his own notion of integrability, which is so strong it goes well beyond what is now called 'complete' integrability: it involves confining the motion to a one-dimensional *closed* manifold—"die Bahn ist dann eine geschlossene, ihre Punkte bilden ein Kontinuum von nur einer Dimension."¹¹ He seems to reduce integrability

⁵Even if it would make no *invariant* sense, no *geometrical* sense, some unfortunate *coordinate-dependent* sense could perhaps be salvaged.

⁶Symplectically at any rate; certainly compared to Whittaker (1917), for instance.

⁷Einstein (1917a) p. 90

⁸Einstein (1917a,b,c)

⁹Even today, *geometric* mechanics can be symplectic or Riemannian; see Calin & Chang (2005), Pettini (2007). Needless to say, Riemannian geometry is much closer to the curved Lorentzian geometry Einstein had practically invented and was thoroughly used to.

¹⁰In §5 below there will be the conserved quantities $f = (f_1, \dots, f_l) \in \mathbb{R}^l$.

¹¹Einstein (1917a) p. 88

and separability¹² (§6 below) to two simple ‘topological’ distinctions: whether the motion eventually closes, or never does (which would pose an *irremediable* problem for quantisation); and if it does close, whether it intersects itself (which poses a *remediable*, indeed most welcome, problem for quantisation—so welcome that Einstein’s whole invariant, topological agenda rests on it). Again, self-intersections are remedied by *Riemannisierung*, in other words enlargement of the configuration space, considered in the next section.

2 The two quantum rules

Sommerfeld’s quantum conditions

$$(1) \quad J_i^* = \oint p_i dq_i = n_i h,$$

$i = 1, \dots, l$, rule out atomic motions whose actions J_i^* are not integer multiples $n_i \in \mathbb{Z}$ of Planck’s constant h . Having spent the previous years immersed in the tensorial covariance of general relativity, Einstein sees a problem here: Sommerfeld’s rule refers to the specific coordinates q_1, \dots, q_l ; each one of the l conditions concerns a particular momentum component p_i . Einstein replaces them with l conditions, each one of which involves the *entire* momentum one-form

$$(2) \quad \mathbf{p} = \sum_{i=1}^l p_i dq_i,$$

now integrated over the l homotopy classes $\mathbb{H}_1, \dots, \mathbb{H}_l$ or ‘topological features’ of the space—there being a quantum condition for every such feature. To understand the construction we can begin with an annulus.

Einstein considers the example of an annulus Ω bounded by circles of radii r_1, r_2 . Motions on Ω that never close are *intractable*; those that close without intersecting themselves are *too tractable*—for Sommerfeld’s rule in particular, thereby giving Einstein no advantage. Einstein’s whole strategy relies on closed motions that intersect themselves: the only motions combining the two virtues of being somehow tractable, but not so much as to preclude the welcome difficulties in whose solution lies Einstein’s real edge over Sommerfeld, and which represent the chief interest of *Quantensatz*.

To see the problem, the above ‘topological’ classification of motions can be re-expressed in terms of momentum assignments to (or right around) points: [1] *one*, [2] *finitely many*, [3] *infinitely many*. Einstein considers an infinitesimal region¹³ $\tau \subset \Omega$ crossed by the motion, the following cases can arise:

- [1] The next time the motion crosses τ it assigns to $\mathbf{q} \in \tau$ the same momentum \mathbf{p} —and hence every time thereafter. This is the simplest kind of periodicity: the motion is closed (and hence confined to a loop, a one-dimensional manifold of finite length) and never intersects itself.

¹²See Gutzwiller (1990) §3.7 on the *modern* distinction (not Einstein’s).

¹³To include ergodic motion in the classification Einstein broadens his purview from the point to its immediate surroundings.

- [2] The motion eventually assigns a *finite* number N of momenta¹⁴ $\mathbf{p}_1, \dots, \mathbf{p}_N$ to τ , finally closing on the N -th lap. The *Bahn* is still *eine geschlossene, ihre Punkte bilden ein Kontinuum von nur einer Dimension*; but it intersects itself.¹⁵
- [3] The motion assigns an *infinite* number of momenta to τ , without ever closing. Being ergodic, the motion cannot be confined to a loop, an *exakt geschlossene Bahn*.

The quantisation rules at issue here only make sense with a single momentum at (or right around) a point. To make the momentum assignment amenable to quantisation, Einstein enlarges the configuration space, thus restoring single-valuedness. But the enlargement procedure he adopts is *finite*, and cannot be repeated *infinitely* many times (for infinitely many momentum values). *Riemannisierung* cannot be applied to ergodic motions.

So Einstein takes a closed but nontrivial *Rosettenbahn*, which intersects itself. He chooses a point $\mathbf{q} \in \mathcal{Q}$ to which the motion assigns two momenta, \mathbf{p}_1 and \mathbf{p}_2 . In order to restore the *Einwertigkeit* needed for quantisation he superposes a second annulus on the first, identifying the delimiting circles, and stipulating that whenever the motion reaches either one it changes annulus. The motion on the resulting two-torus \mathfrak{T}^2 can now be quantised since it no longer intersects itself. The topology of the torus is captured by the (nontrivial) homotopy classes¹⁶ \mathbb{H}_1 and \mathbb{H}_2 , respectively made up of loops going around the first and second circles of the torus (once).¹⁷ The integral

$$(3) \quad \langle \mathbf{p}, \mathbb{H}_i \rangle = \oint_{\mathbb{H}_i} \mathbf{p}$$

vanishes for neither \mathbb{H}_1 nor \mathbb{H}_2 , whereas $\langle \mathbf{p}, \mathbb{H}_0 \rangle$ does vanish for the trivial homotopy class \mathbb{H}_0 of ‘contractible’ loops going around neither circle: Einstein specifies that the one-form \mathbf{p} on \mathfrak{T}^2 is κ -closed¹⁸ (or perhaps even exact—see §4 below). The new quantum rule¹⁹

$$(4) \quad \langle \mathbf{p}, \mathbb{H}_i \rangle = n_i h$$

$i = 1, \dots, l$, which in fact applies more generally to any l -torus²⁰ \mathfrak{T}^l , requires the loop integrals to be integer multiples $n_i \in \mathbb{Z}$ of Planck’s constant h . The methodological, æsthetic superiority of the rule lies in the double invariance of the expression

$$\oint_{\mathbb{H}_k} \sum_i p_i dq_i :$$

¹⁴Each $\mathbf{p}_k \in \wedge^1 \mathcal{Q}$ is a one-form on \mathcal{Q} and not a coordinate $\in \mathbb{R}$.

¹⁵See Gutzwiller (1990) figures 10 & 30.

¹⁶Graffi (2005) p. 23: “curve topologicalmente inequivalenti.”

¹⁷See Gutzwiller (1990) figure 30.

¹⁸Since I am using “closed” in two entirely different senses, I’ll write “ κ -closed” for this second sense (‘locally exact’), pertaining to differential forms.

¹⁹See Graffi (2004) p. 175.

²⁰In the paragraph on p. 90 containing figures 1 & 2 Einstein seems to have in mind an l -dimensional *torus* \mathfrak{T}^l (rather than a more general manifold with Betti number l). In two dimensions, *Riemannisierung* clearly produces a torus; little generality is lost in considering \mathfrak{T}^2 ; and one wonders how the scheme can work in general if the enlarged configuration space is not toroidal. So I’ll speak of a torus \mathfrak{T}^l even with $l > 2$.

invariance with respect to the choice of coordinates q_i , and also with respect to the particular loop of \mathbb{H}_k —with respect to ‘loop deformations.’²¹ All ‘old quantum theories’ are empirically too inadequate to justify an *empirical* preference of Einstein’s rule over Sommerfeld’s.

Again, both quantum rules work in case [1], neither one in cases [2] and [3]; but at least in case [2], the multi-valuedness (being finite) can be resolved on a larger configuration space \mathfrak{X}^l , where Einstein’s rule works well and makes perfect sense. Even Sommerfeld’s conditions can be *made to* work on \mathfrak{X}^l —which, however, strongly favours the double invariance of Einstein’s rule.

3 Loops, *trajectoires*, gauge

Einstein’s loops are meaningless on their own and only acquire significance collectively as elements of the homotopy classes $\mathbb{H}_1, \dots, \mathbb{H}_l$ which capture the topological peculiarities of the enlarged configuration space. Even if Einstein goes to the trouble of writing “irgendeine geschlossene Kurve, welche durchaus keine „Bahnkurve“ des mechanischen Systems zu sein braucht,” Broglie (1924) p. 63 nonetheless speaks of “*trajectoires fermées*.” But Broglie’s misunderstanding is in fact more interesting than one may imagine, interesting enough to deserve a few words.

Consider such a mechanical *trajectoire*²² ξ in the plane \mathbb{P} , and the action integral

$$(5) \quad J^* = \oint_{\xi} (p_1 dq_1 + p_2 dq_2)$$

calculated with respect to the coordinates q_1, q_2 . To simplify we can confine the ‘source’²³ to the origin $\mathbf{0} \in \mathbb{P}$; in other words the curl $d\mathbf{p}$ of the momentum one-form

$$\mathbf{p} = p_1 dq_1 + p_2 dq_2$$

vanishes everywhere else, on all of $\bar{\mathbb{P}} = \mathbb{P} - \{\mathbf{0}\}$. The integral (5) has various interesting symmetries²⁴ which are worth looking at. A diffeomorphism

$$\gamma : \mathbb{P} \rightarrow \mathbb{P}; \quad \mathbf{q} \mapsto \mathbf{Q} = \gamma(\mathbf{q}); \quad \xi \mapsto \Xi = \gamma(\xi)$$

defined on all of \mathbb{P} displaces everything—*trajectoire*, source at the origin, and other points $\bar{\mathbb{P}}$ —so as to preserve the relations of inclusion and exclusion, and hence the integral J^* itself. Even a diffeomorphism $\bar{\gamma} : \bar{\mathbb{P}} \rightarrow \bar{\mathbb{P}}$ on $\bar{\mathbb{P}}$ (as opposed to \mathbb{P})—which only displaces points where $d\mathbf{p}$ vanishes, and the *trajectoire* itself, but not the source $\mathbf{0}$ —would be just as symmetric, as it could never drag ξ over the source. For a diffeomorphism to alter the topological relations on which (5) depends it would have to displace *selectively*, telling apart origin and points of ξ .

²¹Cf. Graffi (2005) p. 22.

²²If the Hamiltonian has no explicit dependence on time, the *trajectoire* is best viewed as a mere (one-dimensional) manifold, since the progression of its (temporal) parameter would be trivial; see §4 below.

²³The term “source” comes from the ‘divergence’ version of Stokes’s theorem (attributed to Gauß or Ostrogradsky or perhaps Green), here it is more metaphorical. Here the ‘source’ at the origin produces a turbulence which by the theorem manifests itself on the loop ξ as the circulation J^* .

²⁴See Epstein (1916a) p. 172, Whittaker (1917) pp. 271ff, Levi-Civita & Amaldi (1974) pp. 353ff.

Since γ pulls a real-valued function

$$Q_i : \mathbb{P} \rightarrow \mathbb{R}; \mathbf{Q} \mapsto Q_i(\mathbf{Q})$$

defined on the range \mathbb{P} back²⁵ to the domain \mathbb{P} of γ , yielding a function

$$q_i = \gamma^{\otimes} Q_i = Q_i \circ \gamma : \mathbb{P} \rightarrow \mathbb{R}; \mathbf{q} \mapsto q_i(\mathbf{q}) = (\gamma^{\otimes} Q_i)(\mathbf{q})$$

on the domain \mathbb{P} , it pulls the differential dQ_i back accordingly—and hence the basis dQ_1, dQ_2 , and with it the linear combination:

$$\mathbf{p} = \gamma^{\otimes} \mathbf{P} = p_1 \gamma^{\otimes} dQ_1 + p_2 \gamma^{\otimes} dQ_2,$$

where I have written \otimes to avoid confusion with the asterisk Einstein uses to denote the Maupertuis action J^* . We can write

$$J^* = \oint_{\xi} \mathbf{p} = \oint_{\Xi} \mathbf{P},$$

or even

$$J^* = \oint_{\Xi} \mathbf{p} = \oint_{\xi} \mathbf{P},$$

and

$$\begin{aligned} (6) \quad J^* &= \oint_{\gamma'(\xi)} \gamma^{\otimes} \mathbf{p} \\ &= \oint_{\mathbb{H}} \gamma^{\otimes} \mathbf{p} = \langle \gamma^{\otimes} \mathbf{p}, \mathbb{H} \rangle \end{aligned}$$

for all diffeomorphisms

$$\gamma, \gamma' : \mathbb{P} \rightarrow \mathbb{P}$$

leaving the source inside the *trajectoire*. Broglie's *trajectoires* thus become highly transformable, invariant entities—little more than *deformable loops expressive of topological properties*. It remains a mistake to think of Einstein's loops $\xi \in \mathbb{H}$ as *trajectoires*, but not an uninteresting one.

In Hamiltonian mechanics one distinguishes between *point* transformations γ (on the l -dimensional q_i -*Raum*) and *canonical* transformations (on the $2l$ -dimensional phase space Γ).²⁶ So far we haven't gone beyond the rather restrictive 'point' condition

$$(7) \quad \mathbf{p} = \sum_i p_i dq_i = \sum_i P_i dQ_i$$

of Lagrangian mechanics (where the configuration space \mathcal{Q} characterised by q_1, \dots, q_l maintains its identity, without 'getting lost' in the $2l$ -dimensional state space); but in

²⁵The case is so trivial (domain and range coincide *etc.*) that pulling back and pushing forward can be legitimately confused; so the abuse of notation $\gamma^{\otimes} \mathbf{p}$, for instance, is venial. A diffeomorphism can be taken to map from the domain \mathbb{P} to the range \mathbb{P} , or the other way around.

²⁶See Landau & Lifschitz (1970) § 45.

fact the weaker ‘symplectic’ condition

$$(8) \quad \begin{aligned} \omega &= d\mathbf{p} = d\mathbf{p}' = d(\mathbf{p} + dF) \\ &= \sum_i dp_i \wedge dq_i = \sum_i dp_i \wedge dq_i + d^2F \end{aligned}$$

also preserves J^* , where

$$(9) \quad \mathbf{p} \mapsto \mathbf{p}' = \mathbf{p} + dF$$

and the generating function F is a zero-form. So in addition to the diffeomorphic freedom γ (or $\bar{\gamma}$) and the homotopic freedom $\xi \mapsto \Xi$ (keeping clear of the source), we have the ‘gauge’ or ‘symplectic’ freedom (8) to add an exact term dF :

$$J^* = \langle \gamma^{\otimes} \mathbf{p} + dF, \mathbb{H} \rangle.$$

While a diffeomorphism γ affects the momentum one-form indirectly by first dragging points, the gauge transformation (9) is fibre-preserving and therefore acts directly on each $\mathbf{p}(\mathbf{q})$, point by point. The canonical transformation generated by F also corresponds to a deformation of curves in $\bar{\mathbb{P}}$, reminiscent of the loop deformation $\xi \mapsto \Xi$. Even if the curl at the origin (or the corresponding circulation J^*) prevents the one-form \mathbf{p} from having a *global* primitive, one can think of a *local* primitive λ satisfying $\mathbf{p} = d\lambda$ locally.²⁷ In much the same way as one can take ξ to be displaced by the point transformation γ , one can take the level sets $\lambda = \text{const.}$ to be deformed by F .

Graffi (2005) has rightly pointed out that Einstein’s rule (4) is *canonically* invariant. The symplectic freedom (8) is undeniably *available*; but Einstein’s mechanics is so *unsymplectic*, so firmly rooted in the q_i -*Raum* he so often mentions, that I doubt he had in mind anything beyond the two genuinely q_i -*Raum* symmetries represented in (6).

4 Hamilton-Jacobi-Einstein

The old quantum theory, including *Quantensatz*, was formulated in terms of Hamilton-Jacobi theory,²⁸ which therefore deserves some attention.

The principle of least action²⁹ exists in two versions, which in Einstein’s notation would be distinguished by a star indicating neglect of time.

[i] The ‘space-time’ Hamiltonian version determines the spatial shape of trajectories *as well as motion along them* by minimising the full Hamiltonian action³⁰

$$J = \int_{t_0}^t \mathcal{L} dt,$$

²⁷“Locally” could mean, for instance, *on any simply-connected region of $\bar{\mathbb{P}}$* .

²⁸See Hamilton (1833, 1834), Jacobi (1884) pp. 143ff, Appell (1909) pp. 556ff, Epstein (1916b) pp. 493ff, Whittaker (1917) pp. 288ff, Carathéodory (1937) pp. 66ff, Brillouin (1938) pp. 168ff, Appell (1953) pp. 429ff, Levi-Civita & Amaldi (1974) pp. 355ff, Arnol’d (1988) §9.4, Jacobi (1996) pp. 216ff, Fasano & Marmi (2002) §11.1, Graffi (2004) p. 51, Benci & Fortunato (2014) §1.4.

²⁹See Brillouin (1938) pp. 159ff.

³⁰*Cf.* Carathéodory (1937) p. 10.

where the ‘momentum’ or ‘covariant’ Lagrangian can be written

$$\mathcal{L} = \sum_i p_i \frac{\partial \mathcal{H}}{\partial p_i} - \mathcal{H}.$$

[ii] The ‘spatial’ version attributed to Maupertuis only gives the spatial shape of the trajectory, by minimising the purely spatial part

$$(10) \quad J^* = \int_{\mathbf{q}_0}^{\mathbf{q}} \mathbf{p} = \int_{\mathbf{q}_0}^{\mathbf{q}} \sum_i p_i dq_i$$

of the action.

Hamilton’s principle is in a sense more general; but the generality it adds to Maupertuis’ is only of any interest if the Hamiltonian \mathcal{H} depends explicitly on time. Since Einstein takes it not to, the motion is confined to a level surface where the Hamiltonian remains equal to some constant E ; and the action assumes the degenerate additive form³¹

$$J = J^* - Et.$$

Let us write $\mathcal{H} = T + U$ (and $\mathcal{L} = T - U$), where the potential U depends only on position and the kinetic energy

$$\begin{aligned} T &= \frac{\|\mathbf{p}\|^2}{2m} = \frac{1}{2} \sum_i p_i \dot{q}_i \\ &= \frac{1}{2} \sum_i p_i \frac{\partial \mathcal{H}}{\partial p_i} \end{aligned}$$

is quadratic in the momenta. Only the spatial shape of the trajectory would then remain interesting. To confine our attention to that shape, ignoring the trivial time evolution given by the term $-Et$, we just take the spatial ‘Maupertuis’ part of J , namely (10). Viewing \mathbf{q}_0 as a fixed initial position and \mathbf{q} as a variable final position, we can radiate \mathbf{q} in all directions from \mathbf{q}_0 (now reminiscent of a source in geometrical optics³²) along dynamical trajectories derived from the action function $J^*(\mathbf{q}, \mathbf{q}_0)$ satisfying the Hamilton-Jacobi equation³³

$$(11) \quad \mathcal{H}(\mathbf{q}, dJ^*) = E,$$

or rather

$$\|\mathbf{p}\|^2 = \|dJ^*\|^2 = 2m(E - U).$$

Choosing an infinitesimal action δJ^* , we first have the sphere $\sigma(\mathbf{q}_0, \delta J^*)$ of radius

$$\zeta(\mathbf{q}_0, \delta J^*) = \frac{\delta J^*}{\sqrt{2m(E - U(\mathbf{q}_0))}}$$

³¹See Jacobi (1884) *Einundzwanzigste Vorlesung*, Appell (1953) pp. 430ff, Levi-Civita & Amaldi (1974) pp. 362ff.

³²See Hamilton (1833, 1834), Whittaker (1917) pp. 288ff.

³³See Graffi (2004) p. 51.

around \mathbf{q}_0 . The momentum \mathbf{p} at any \mathbf{q} on the sphere is the differential

$$\mathbf{p}(\mathbf{q}) = dJ^*(\mathbf{q}) = \sum_{i=1}^l \frac{\partial J^*(\mathbf{q})}{\partial q_i} dq_i$$

of J^* viewed as a function of the final position \mathbf{q} . Once we have a level surface $\sigma(\mathbf{q}_0, \delta J^*)$ of J^* , we might as well generalise and propagate from an *arbitrary* $(l-1)$ -dimensional initial surface

$$\Sigma_1 = \Sigma_{J^*_{[1]}}$$

(which could be a sphere σ or not), viewed as a level surface of action $J^*_{[1]}$. I'll follow Schrödinger (1926, pp. 492-3) on the way to his *Wellenmechanik*.³⁴ We can again take the same increment³⁵ δJ^* , which provides a distance $\zeta(\mathbf{q}_1, \delta J^*)$ at the generic point $\mathbf{q}_1 \in \Sigma_1$. Then there are two possible constructions: Either we repeat the above construction, now treating \mathbf{q}_1 as the source of a 'secondary' wave,³⁶ a sphere $\sigma(\mathbf{q}_1, \delta J^*)$ of radius $\zeta(\mathbf{q}_1, \delta J^*)$ emanated from every point of the 'primary' wavefront Σ_1 , in which case the surface Σ_2 of action

$$J^*_{[2]} = J^*_{[1]} + \delta J^*$$

is the envelope of the secondary waves. Alternatively one lays off $\zeta(\mathbf{q}_1, \delta J^*)$ normally at every point, to define Σ_2 even more directly. 'Normally' can be understood as follows: The $(l-1)$ -dimensional linear space

$$\mathbb{T}_{\mathbf{q}_1} \Sigma_1 \subset \mathbb{T}_{\mathbf{q}_1} \mathcal{Q}$$

tangent to Σ_1 at \mathbf{q}_1 determines a ray ρ_1^\flat in the cotangent space $\mathbb{T}_{\mathbf{q}_1}^* \mathcal{Q}$. The direction of the momentum \mathbf{p}_1 at \mathbf{q}_1 is given by $\mathbf{p}_1 \in \rho_1^\flat$, the length by

$$\|\mathbf{p}_1\| = \sqrt{2m(E - U(\mathbf{q}_1))}.$$

The inverse m^\sharp of the mass tensor³⁷ m^\flat determines a vector $\mathbf{p}_1^\sharp = m^\sharp \mathbf{p}_1$, indeed a ray ρ_1^\sharp containing \mathbf{p}_1^\sharp in the tangent space $\mathbb{T}_{\mathbf{q}_1} \mathcal{Q}$. The action increment δJ^* fixes the length $\zeta(\mathbf{q}_1, \delta J^*)$ of the required vector $\hat{\mathbf{p}}_1^\sharp \in \rho_1^\sharp$, which goes from Σ_1 to the corresponding point of Σ_2 .

So that's one way (specificially a rather *Hamiltonian* way) of understanding the Hamilton-Jacobi equation (11)—which is so central to the old quantum theory, and even to wave mechanics: as an infinitesimal condition governing the orthogonal propagation in configuration space \mathcal{Q} of a solution J^* from an initial surface (which could be as small as a point). But the orthogonal construction also serves to elucidate, by opposition, the somewhat different logic of the 'oblique' construction Einstein proposes in § 3 of *Quantensatz*. In Schrödinger's construction, the direction of propagation

³⁴See also Brillouin (1938) pp. 169ff, Arnol'd (1988) pp. 251ff.

³⁵A smaller increment would be even better, if secondary waves are propagated from an *already infinitesimal* primary wavefront $\sigma(\mathbf{q}_0, \delta J^*)$.

³⁶See Huygens (1690), Whittaker (1917) pp. 289ff, Carathéodory (1937) pp. 13ff, Arnol'd (1988) §9.3.

³⁷See Brillouin (1938) p. 143.

from the initial surface Σ_1 is determined by the local slant of the surface itself,³⁸ not by initial data freely assigned to it, point by point. Einstein also propagates from an $(l - 1)$ -dimensional surface in the q_i -Raum \mathcal{Q} , but *obliquely*—at an ‘angle,’ relative to the surface, that varies from point to point. Schrödinger’s orthogonal propagation, which makes the initial surface Σ_1 a *level* surface of J^* , also makes the propagated field *exact*, indeed consistent with a Hamilton-Jacobi potential J^* ; Einstein’s ‘oblique’ construction has to require exactness independently, as it is not provided by orthogonality.

The difference between κ losed and *exact*, it should be noted, may not have been entirely clear to Einstein (or others in 1917). Important progress is made in Rham’s *Thèse* (1931), where the distinction is used for *analysis situs*. Even if the distinction is clear to Weyl (1939), he confusingly uses the word “exact” to mean κ losed: “A differential ω whose derivative vanishes is called *exact*.” Weyl’s understanding of κ losed is already rather modern:

What we mean by $\omega \sim 0$ (ω homologous zero) may be explained in two ways: either differentially as indicating that ω is the derivative of a differential of next lower rank, or integrally as demanding that the integral of ω over any cycle vanishes. Every differential ~ 0 is exact; one readily proves this in both ways. “In the small” both notions, exact and ~ 0 , coincide, but not in the large.

But one can imagine how roughly, if at all, the distinction was understood in 1917. In (1917c) Einstein seems to treat his equations (7) and (7a) as equivalent; in *Quantensatz* he writes “bzw.” between his equations 10) and 10a)—but then does draw a related distinction, between *einwertig* and *vielwertig*, on the next page:

Ist aber der in Betracht kommende Raum der q_i ein mehrfach zusammenhängender, so gibt es geschlossene Bahnen, welche nicht durch stetige Änderung auf einen Punkt zusammengezogen werden können; ist dann J^* keine einwertige (sondern eine ∞ vielwertige) Funktion der q_i , so wird das Integral $[\int \sum p_i dq_i]$ für eine solche Kurve im allgemeinen von Null verschieden sein.

Indeed if the one-form \mathbf{p} is only

CMC: *closed on a multiply-connected region,*

it may or may not be exact, the loop integral (3) may or may not vanish. If \mathbf{p} is not exact,³⁹ the loop integral $\langle \mathbf{p}, \mathbb{H} \rangle$ will not vanish, but its *Vielwertigkeit* will be the *denkbar einfachste*: every point \mathbf{q} on a given loop acquires an additional $\langle \mathbf{p}, \mathbb{H} \rangle$ on every lap:

$$\begin{aligned} J_{(n+1)}^*(\mathbf{q}) &= J_{(n)}^*(\mathbf{q}) + \langle \mathbf{p}, \mathbb{H} \rangle \\ &= J_{(1)}^*(\mathbf{q}) + n \langle \mathbf{p}, \mathbb{H} \rangle, \end{aligned}$$

³⁸Here the normal direction of propagation is determined by the mass tensor m . With a more general Hamiltonian (and Lagrangian), the duality relations $\sharp \flat$ between level surface and motions are given by the fibre derivatives, with components $\partial \mathcal{H} / \partial p_i, \partial \mathcal{L} / \partial q_i$.

³⁹The ‘topological interpretation’ of the Aharonov-Bohm effect shows how important it is to distinguish between *not* exact and the above condition **CMC**: not *necessarily* exact.

where the integer n stands for the lap. In terms of the *Ein/Vielwertigkeit* of the primitive J^* , *exact* means *has a single-valued global primitive J^** ; merely κ losed means *exact “in the small”*—the primitive J^* is *locally einwertig* but may be globally *vielwertig*.

Returning to Einstein’s construction, he freely assigns momenta $\mathbf{p}_1(\mathbf{q}_1)$ to the points \mathbf{q}_1 of Σ_1 , then radiates dynamical (Hamiltonian) trajectories⁴⁰ L satisfying

$$\dot{L}(\mathbf{q}_1) = \mathbf{p}_1^\sharp(\mathbf{q}_1) = m^\sharp \mathbf{p}_1(\mathbf{q}_1)$$

throughout Σ_1 . The dynamical vector field \dot{L} thus determined by the one-form $\mathbf{p}_1 = \mathbf{p}|_{\Sigma_1}$ provides a one-form $\mathbf{p} = m^\flat \dot{L}$ on an l -dimensional region $\mathcal{U} \subset \mathcal{Q}$ covered by \dot{L} .

Einstein then wonders when the one-form \mathbf{p} determined by the Hamiltonian vector field \dot{L} will also satisfy the Hamilton-Jacobi equation; for the radiated congruence can, confusingly, be ‘Hamiltonian’ (in other words *made up of dynamical trajectories*) without being ‘Hamilton-Jacobi’—consistent with a potential J^* satisfying the Hamilton-Jacobi equation. It turns out that \mathbf{p}_1 has to be κ losed, for then \mathbf{p} will be too; and as long as \mathcal{U} is topologically simple, κ losed means exact:

$$(d\mathbf{p}_1 = 0) \Leftrightarrow (d\mathbf{p} = 0) \Leftrightarrow (\exists J^* : \mathbf{p} = dJ^*).$$

The potential J^* is slightly underdetermined by its derivative dJ^* , in other words

$$d^{-1}\mathbf{p} = [J^*] = [J^* + \eta]_\eta,$$

where η is a constant on \mathcal{U} . A value $J^*(\mathbf{q})$ at a single $\mathbf{q} \in \mathcal{U}$ is enough to overcome the underdetermination and fix all of J^* . Summing up, the theorem can be given as follows: Only a κ losed momentum field \mathbf{p} (radiated *dynamically* from the momenta \mathbf{p}_1 freely assigned to an initial surface $\Sigma_1 \subset \mathcal{Q}$) can be derived from a potential

$$J^* = d^{-1}\mathbf{p} - \eta$$

satisfying the Hamilton-Jacobi equation.

Einstein seems to change his mind in the *Nachtrag zur Korrektur*, p. 91:

LB: Liefert eine Bewegung ein p_i -Feld, so besitzt dieses notwendig ein Potential J^* .

Whatever he means has to do with § 4 of *Quantensatz*: “die zweite der in § 4 angegebenen Bedingungen für die Anwendbarkeit der Formel 11) stets von selbst erfüllt sein muß [...]”; at the end of § 4 we discover that

Die Anwendung der Quantenbedingung 11) verlangt, daß derartige Bahnen existieren, daß die einzelnen Bahn ein p_i -Feld bestimmt, für welches ein Potential J^* existiert.

⁴⁰The almost superfluous curves L and vectors \dot{L} used to extend the one-form \mathbf{p}_0 on Σ_0 to a one-form \mathbf{p} on the l -dimensional region \mathcal{U} are only introduced because Einstein seems to have such objects in mind. They can easily be dispensed with, the one-forms are enough.

To try to make sense of this we can start with the primitive notion of *Bewegung*: a motion, a trajectory in configuration space—however it may be defined or generated. Returning to the above classification, the motion can be closed (case [1] or [2]) or ergodic (case [3]). If it is closed, it will assign to no point \mathbf{q} (or neighbourhood) of \mathcal{Q} more than finitely many momenta $\mathbf{p}_1, \dots, \mathbf{p}_N$. That’s what Einstein means by a p_i -Feld: an assignment of at most finitely many momenta to (certain) points of \mathcal{Q} . If the momentum is *finitely* multi-valued here and there, the configuration space can be enlarged to restore single-valuedness—but not if it is ergodic.

There remains the issue of how a single *Bewegung*, confined as it is to a one-dimensional manifold, can yield a *Feld*, a field on an l -dimensional manifold. Since Einstein goes to the trouble of emphasising *die einzelne Bahn* with *Sperdrück*, he really does seem to mean a *single* trajectory. He may simply be unaware of the problem; or perhaps it somehow doesn’t bother him and he just ignores it; or perhaps he has a way of dealing with it. But if he does, it would surely have to involve a *congruence*, a *Schar*, somehow or other—despite the arresting *Sperdrück* of *die einzelne Bahn*.

Indeed Einstein’s use of Hamilton-Jacobi theory is most peculiar, perhaps even contradictory or downright wrong. It is a theory that undeniably involves *congruences* of trajectories. In § 3 of *Quantensatz* and in (1917c), Einstein considers the entire congruence; but elsewhere in *Quantensatz* he seems to use the same theory to produce a *single* trajectory. The integral (3) only makes invariant sense if \mathbf{p} is defined *everywhere* on \mathcal{Q} , not just on a single trajectory—Einstein accordingly speaks of an entire p_i -Feld. But Einstein’s whole analysis and classification of trajectories (our cases [1]-[3])—which would get confused, perhaps even undermined, by congruences—seems to depend on *single* trajectories. We may simply have a case of Einstein wanting to have his cake and eat it.

Returning to **LB**, the issue of a Hamilton-Jacobi potential J^* or of a vanishing curl $d\mathbf{p}$ does not even arise with a single trajectory. With a whole congruence of motions radiated from an initial surface, the curl $d\mathbf{p}$ automatically vanishes if the propagation is orthogonal, as in Schrödinger’s construction. If the propagation is only transversal, as in § 3 of *Quantensatz*, the curl has to vanish for the motions to admit a Hamilton-Jacobi potential J^* .

5 Liouville-Einstein-Arnol’d?

In the second-last paragraph of the *Nachtrag* Einstein formulates an integrability theorem that deserves attention:

IT \mathcal{Q} : Existieren l Integrale der $2l$ Bewegungsgleichungen von der Form

$$(12) \quad R_k(q_i, p_i) = \text{konst.},$$

wobei die R_k algebraische Funktionen der p_i sind, so ist $\sum_i p_i dq$ immer ein vollständiges Differential, wenn man die p_i vermöge (12) durch die q_i ausgedrückt denkt.

The formulation is misleading, he may be groping towards the theorem now attributed to Liouville⁴¹ and Arnol'd.⁴²

We can begin anachronistically with (features of) the Liouville-Arnol'd theorem, and then try to understand how much of it is already in *Quantensatz*. The theorem uses l functions

$$F_i : \Gamma \rightarrow \mathbb{R}$$

on an $2l$ -dimensional phase space Γ to reduce the number of dynamically relevant dimensions from $2l$ to l ; in the sense that by $l - 1$ intersections of level surfaces of appropriately compatible and independent functions it confines the dynamics to an l -dimensional manifold. A function F foliates the phase space into $(2l - 1)$ -dimensional level surfaces on which $F = \text{const.}$; specifying a value f of F already eliminates one dimension by fixing a level surface. But the theorem concerns dynamics; the issue is whether a given dynamics⁴³

$$X_G = (dG)^\sharp = \omega^\sharp(dG)$$

is *tangent* to the level surfaces $F = \text{const.}$, or transversal to them. If the dynamics X_G were transversal to the level surfaces of F , the dimension lost by choosing a level surface would be thus restored, with no net progress in the effort to eliminate dimensions.

The relevant notion is Poisson compatibility⁴⁴

$$\{F, G\} = 0,$$

which can either be understood as compatibility between the dynamics X_F generated by F and the level surfaces of G , or the other way around. $\{F, G\}$ vanishes if the vector field X_F is tangent to the level surfaces of G , in other words if the graph of each integral curve of X_F is confined to a level surface of G . We can take the first function F_1 to be the Hamiltonian \mathcal{H} , whose generic value E singles out a $(2l - 1)$ -dimensional energy surface. The two compatibilities

$$\{\mathcal{H}, F_2\} = 0 = \{\mathcal{H}, F_3\}$$

and values $F_2 = f_2, F_3 = f_3$ only eliminate both dimensions (and not just one) if dF_2 and dF_3 are independent; if dF_2 and dF_3 were parallel, F_2 and F_3 would have the same level surfaces, which would be redundant.⁴⁵ *Complete integrability* is given by l values

$$f = (f_1, \dots, f_l)$$

of the compatible, independent functions F_1, \dots, F_l , which eliminate l of the initial $2l$ dimensions of phase space, leaving an l -dimensional manifold \mathcal{M}_f . Einstein's integrability theorem **IT**_Q suggests he may have got this far. But the two cases at the bottom

⁴¹See Appell (1909) pp. 576ff, Appell (1953) pp. 437ff, Fasano & Marmi (2002) §11.4.

⁴²See Fasano & Marmi (2002) §11.5, Graffi (2004) §1.7.1, Graffi (2005) §2, Lowenstein (2012) pp. 56ff.

⁴³See Arnol'd (1988) §8.1.3, Sternberg & Guillemin (1984) p. 88, Graffi (2004) p. 51.

⁴⁴See Graffi (2004) p. 50.

⁴⁵In the aforementioned 'effort to eliminate dimensions,' Poisson incompatibility $\{F, G\} \neq 0$ would be *counterproductive* (reversing progress already made by effectively *restoring* an eliminated dimension), whereas linear dependence $dF = kdG$ would be merely *unproductive* (producing neither loss nor gain, leaving the number of dimensions unchanged).

of p. 87 of *Quantensatz* give the impression that Einstein wants to reduce the dynamically available dimensions of q_i -Raum, not phase space; that he wants to confine the *Bahnkurve* to a lower-dimensional manifold of *configuration space*:

Die Bahnkurve läßt sich ganz in einem Kontinuum von weniger als l Dimensionen unterbringen.

The kind of integrability he has in mind, taken as far as possible, would confine the *Bahnkurve* to a *one-dimensional* submanifold of Q , not Γ :

Hierzu gehört als spezieller Fall derjenige der Bewegung in exakt geschlossener Bahn. [...] die Bahn ist dann eine geschlossene, ihre Punkte bilden ein Kontinuum von nur einer Dimension.

This brings us to the next part of the Liouville-Arnol'd theorem. So far we have little more than a number l of dimensions. Without compactness, the l -dimensional manifold \mathcal{M}_f could be a product of lines and loops; compactness rules out the lines, leaving a torus \mathfrak{T}^l , a product of l topological circles. It is worth noting that Arnol'd and Einstein obtain their tori in different ways: Arnol'd by imposing compactness, Einstein by *Riemannisierung*, to eliminate self-intersections. The two ways, however different, are not unrelated: if the l -dimensional manifold \mathcal{M}_f were a product of l lines, it would amount to \mathbb{R}^l ; compactness prevents immersion in \mathbb{R}^l by 'swelling' \mathcal{M}_f . A two-dimensional torus \mathfrak{T}^2 , for instance, is an *enlarged* two-dimensional configuration space inasmuch as it cannot be embedded in the plane. Suppose for definiteness that \mathfrak{T}^2 is not just a 'topological torus' (a product of two loops) but a 'rigid torus' (a product of two rigid circles \mathbb{S}^1)—literally a doughnut, with a shape and a size, embedded in \mathbb{R}^3 parallel to the xy plane. Take the simplest possible motion, given by fixed rates of rotation around both circles of such a \mathfrak{T}^2 : projected onto the xy plane it would intersect itself (at regular intervals). That's how compactness is related, albeit indirectly, to Einstein's *Riemannisierung*.

The distinction between rigid and topological tori, introduced above for mere definiteness of representation, will actually prove quite relevant. According to the Liouville-Arnol'd theorem, a completely integrable (and compact) dynamics can be represented as l constant rates

$$(13) \quad \dot{\varphi}_i = \omega_i = \frac{\partial \mathcal{H}}{\partial J_i^*}$$

of rotation on an l -dimensional torus \mathfrak{T}^l , where the l angles⁴⁶ φ_i with linear evolutions

$$\varphi_i(t) = \omega_i t + \varphi_i(0)$$

are canonically conjugate

$$(14) \quad \{\varphi_i, J_j^*\} = \delta_{ij}$$

to the actions J_i^* ; $i, j = 1, \dots, l$.

⁴⁶See Sternberg & Guillemin (1984) pp. 356ff, Lowenstein (2012) pp. 68ff.

Einstein seems to construct a *topological*, l -dimensional torus (topologically equivalent to a rigid torus) to resolve self-intersections; his l integrals (3) are undeniably *action* integrals; but that's not enough to provide the 'symplectic' rigidity that turns a topological torus into a rigid one. To do so, Einstein would have needed something along the lines of (13) or (14)—which were by no means obvious in 1917, especially to a physicist, and cannot be taken for granted. To obtain a rigid torus one has to recognise that the angles φ_i yielding the constant frequencies ω_i are canonically conjugate to the actions J_i^* ; but nowhere does Einstein betray such symplectic awareness. The rigid torus of the Liouville-Arnol'd theorem is really quite different from Einstein's merely topological torus: it has a metrically definite shape and is twice as big, with $2l$ degrees freedom (l areas and l angles), not just l (l arbitrary parameters along l loops). Even if steady motions on a rigid torus are ultimately needed to make sense of his intuitive differentiation of ergodic and closed motions, that distinction alone hardly warrants the attribution of so much definite structure to Einstein's loose, highly topological constructions.

But let us nonetheless consider a rigid torus, which with little loss of generality can be taken to be two-dimensional, with frequencies ω_1, ω_2 . The motion can in any case be confined to a one-dimensional manifold; the issue is its length (finite or not), its topology (closed or open)—whether we have a *Bewegung in exakt geschlossener Bahn*, whose *Punkte bilden ein Kontinuum von nur einer Dimension*. The relevant criterion is rational dependence: if ω_1/ω_2 is rational, the motion can be confined to a one-dimensional *Kontinuum* which, being *geschlossen*, is of finite length; whereas if ω_1/ω_2 is irrational the motion can still be confined to a one-dimensional *Kontinuum*, but not of finite length.⁴⁷ Einstein clearly understands the geometrical significance of confining motion to intersections of level surfaces of appropriately compatible and independent functions; what may be less clear to him is the *numerical*—rational *vs.* irrational—rather than set-theoretical (or 'manifold-theoretical') character of the last step, needed to bring the (appropriately finite) dimensions down to *one*. Indeed in the very last paragraph of the *Nachtrag* he seems to suggest that ergodic motion is only possible with an *incomplete* set R_1, \dots, R_j ($j < l$). We now know that ergodicity can be a merely numerical matter, of *irrational dependence*—compatible with a 'complete' set of l functions R_1, \dots, R_l .

Returning to Einstein's integrability theorem IT_Q , he mentions that the momentum (2) would be a *vollständiges Differential*. I think the point is not exact *vs.* κ losed, but that once ergodic motion is ruled out, with its 'infinite-valued' momentum, the momentum just *makes sense*, being single-valued on an appropriate configuration space. He seems to understand the theorem as follows: his l conserved quantities R_1, \dots, R_l are numerous enough to provide a well-defined momentum \mathbf{p} by confining the motion to an *exakt geschlossene Bahn* whose *Punkte bilden ein Kontinuum von nur einer Dimension*, thus ruling out ergodic motion which, being 'infinite-valued,' is unmanageable.

Einstein's letter (1917b) to Ehrenfest puts it all pretty clearly:

IT_E : Es liege ein Problem vor, bei dem soviel Integrale

$$L(q_\nu, p_\nu) = \text{konst}$$

⁴⁷See Born (1933) pp. 80ff, Fasano & Marmi (2002) §11.7.

existieren, als Freiheitsgrade. Dann können die Impulse als (mehrwertige) Funktionen der q_ν ausgedrückt werden. Andererseits erfülle die Bahnkurve einen gewissen q_ν -Raum vollständig, sodass sie jeden Punkt desselben beliebig nahe kommt. Dann liefert die Bahn des Systems im q_ν -Raum ein Vektorfeld der p_ν .

Andererseits refers to the number of integrals: *if there aren't enough conserved quantities, the motion is ergodic* (case [3] above). The implication being that if there are enough functions L , we have a *Bewegung in exakt geschlossener Bahn*, whose *Punkte bilden ein Kontinuum von nur einer Dimension* ([1] or [2]); l integrals are enough to provide a finitely-many-valued one-form \mathbf{p} (on configuration space \mathcal{Q}), which can be made single-valued by *Riemannisierung*.

6 Einwertigkeit, separability, quantisation

Einstein's imaginative, idiosyncratic treatment of integrability and ergodicity is of the greatest interest; as is the torus he constructs to restore *Einwertigkeit*, not to mention what he puts on it. For us the point is more what Einstein does to restore single-valuedness, than *why* single-valuedness should ever be important in the first place; but a few words on the matter may nonetheless not go amiss.

Even if it could be enough to say that *integrals, like (1) or (3), only make sense for single-valued integrands*, Einstein seems to view it as a matter of *separability*:⁴⁸ quantisation relies on separability, which in turn depends on single-valuedness.

It may be best to think of separability *backwards*, starting with canonical coordinates

$$(J^*, \varphi) = (J_1^*, \dots, J_l^*, \varphi_1, \dots, \varphi_l)$$

which are separable by construction, in the sense that the total action can be written as a sum

$$(15) \quad J^* = \sum_{i=1}^l J_i^*$$

of the l actions $J_i = J_i^*(\mathbb{H}_i)$. The Hamiltonian breaks up accordingly:

$$\mathcal{H}(J_1^*, \dots, J_l^*) = \sum_{i=1}^l \mathcal{H}_i(J_i^*),$$

where the decomposition into l pieces corresponds to the constitution of the torus

$$\mathfrak{T}^l = \mathfrak{T}^l(\mathbb{H}_1, \dots, \mathbb{H}_l) = \prod_{i=1}^l \mathbb{S}_l^1$$

⁴⁸See Appell (1909) pp. 576ff, Epstein (1916a) pp. 170ff, Epstein (1916b) pp. 494ff, Born (1925) §14, Appell (1953) pp. 437ff, Levi-Civita & Amaldi (1974) pp. 405ff, Fasano & Marmi (2002) §11.2, Graffi (2004) p. 175, Lowenstein (2012) pp. 57ff.

itself, a product of l circles \mathbb{S}^1 . A *general* canonical transformation

$$(J_1^*, \dots, J_l^*, \varphi_1, \dots, \varphi_l) \mapsto (p_1, \dots, p_l, q_1, \dots, q_l)$$

from (J^*, φ) cuts across the decomposition by producing new canonical coordinates (p, q) whose lines are *transversal* to those of (J^*, φ) : the entangled general transformation entangles the unentangled coordinates. But the entanglement of coordinates is avoided if the canonical transformation itself is unentangled, in the sense that it breaks up into l canonical transformations

$$(J_1^*, \varphi_1) \mapsto (p_1, q_1), \dots, (J_l^*, \varphi_l) \mapsto (p_l, q_l),$$

one for every part \mathbb{H}_i of the torus, where the symplectic two-form

$$\begin{aligned} \omega &= \sum_{i=1}^l \omega_i = \sum_{i=1}^l dJ_i^* \wedge d\varphi_i = \sum_{i=1}^l d(J_i^* d\varphi_i + dF) \\ &= \sum_{i=1}^l dp_i \wedge dq_i. \end{aligned}$$

If the canonical transformation is generated by an unentangled function

$$F(\varphi_1, \dots, \varphi_l, q_1, \dots, q_l) = \sum_{i=1}^l F_i(\varphi_i, q_i)$$

that decomposes accordingly, the new coordinates (p, q) will be just as separable as (J^*, φ) . Being even-dimensional, each two-dimensional submanifold determined by (J_i^*, φ_i) —by fixing the other $2l - 2$ coordinates of phase space—is a symplectic manifold Γ_i in its own right, with its own symplectic two-form

$$\begin{aligned} \omega_i &= dJ_i^* \wedge d\varphi_i = d(J_i^* d\varphi_i + dF_i) \\ &= dp_i \wedge dq_i \end{aligned}$$

and its own canonical transformations

$$(J_i^*, \varphi_i) \mapsto (p_i, q_i)$$

produced by its own generating functions F_i . If the covector $dF_i(\mathbf{z}) \in \mathbb{T}_{\mathbf{z}}^* \Gamma_i$ is in the coplane

$$\text{span}\{dJ_i^*(\mathbf{z}), d\varphi_i(\mathbf{z})\} \subset \mathbb{T}_{\mathbf{z}}^* \Gamma_i$$

spanned by the covectors $dJ_i^*(\mathbf{z})$ and $d\varphi_i(\mathbf{z})$, it will put both $dp_i(\mathbf{z})$ and $dq_i(\mathbf{z})$ in the same coplane. If that holds at every $\mathbf{z} \in \Gamma_i$, the new pair (p_i, q_i) will determine the same symplectic submanifold Γ_i as the action-angle pair (J_i^*, φ_i) .

That at any rate is a modern notion of separability, which is worth bearing in mind to understand Einstein's. He of course sees things differently. Again, he may not even have the idea of an angle φ_i canonically conjugate to an action J_i^* . He has l actions

$$J_1^*(q_1, \dots, q_l), \dots, J_l^*(q_1, \dots, q_l),$$

or perhaps

$$J_1^*(p_1, \dots, p_l, q_1, \dots, q_l), \dots, J_l^*(p_1, \dots, p_l, q_1, \dots, q_l),$$

where each

$$p_i = p_i(q_1, \dots, q_l)$$

can in general be a function of all l coordinates q_1, \dots, q_l of q_i -Raum. Separability holds when “jedes p_ν nur von dem zugehörigen q_ν abhängt,”⁴⁹ in other words

$$J^* = \sum_{i=1}^l J_i^*(p_i(q_i), q_i).$$

The lines⁵⁰ of the coordinates q_i then turn the Hamilton-Jacobi PDE into l ODEs, each on its own one-dimensional manifold⁵¹ \mathcal{U}_i ; so that the *Unabhängigkeit* on which separability rests is encoded into the very framework $(\mathcal{U}_1, \dots, \mathcal{U}_l)$ of the ODEs. Rather than as a ‘symplectic alignment’ of each (p_i, q_i) along each (J_i^*, φ_i) , Einstein sees separability as a matter of adapting the l coordinates q_1, \dots, q_l of q_i -Raum to the shape of the solution J^* ; as a condition

$$\tau[J^*; (q_1, \dots, q_l)]$$

relating the coordinates—or the *Kontinua* $(\mathcal{U}_1, \dots, \mathcal{U}_l)$ —to the level surfaces of J^* . Locally, around a point $\mathbf{q} \in \mathcal{U}$ in \mathcal{Q} , the coordinates can always be adapted to satisfy the condition: since the infinitesimal behaviour dJ^* of J^* is captured by l derivatives in l independent directions, one can always find l terms J_i^* such that (15) and

$$\frac{dJ_i^*}{dq_i} = \frac{\partial J^*}{\partial q_i}.$$

The real problem is global, not local. As Einstein explains in the *Nachtrag*, J^* may eventually⁵² reach the same \mathbf{q} from different directions; coordinates that satisfy the relationship τ on the first pass may not satisfy it on the next: coordinates can be adapted to a single J^* , but not to two different passes⁵³ $J_{(1)}^*, J_{(2)}^*$. Which is why Einstein has to enlarge the configuration space to resolve self-intersections; in the larger toroidal configuration space, not only are self-intersections eliminated, but separability is restored, since the coordinates can be adapted everywhere to the single-valued trajectory. One can see the problem around a point $\mathbf{q} \in \mathcal{U} \subset \mathcal{Q}$ on the annulus: suppose the self-intersecting motion, the propagation of J^* , first assigns one momentum

$$\mathbf{p}(\mathbf{q}, t_1) = \mathbf{p}_1(\mathbf{q}) = dJ(\mathbf{q}, t_1) = dJ_{(1)}^*(\mathbf{q})$$

⁴⁹Einstein (1917b)

⁵⁰The lines of q_i can be understood as the integral curves of the vector field $\partial/\partial q_i$.

⁵¹In fact \mathcal{U} is partitioned into l independent congruences of one-dimensional manifolds \mathcal{U}_i .

⁵²Even if we are in a configuration space \mathcal{Q} in which time has no more than an *implicit* role, Einstein himself uses explicitly temporal language: “im Laufe der Zeit.”

⁵³In an $(l + 1)$ -dimensional configuration space $\mathcal{Q} \times \mathbb{R}$, enlarged to include time, the two passes would of course take place at different times t_1 and t_2 .

to \mathbf{q} , then a second

$$\mathbf{p}(\mathbf{q}, t_2) = \mathbf{p}_2(\mathbf{q}) = dJ(\mathbf{q}, t_2) = dJ_{(2)}^*(\mathbf{q});$$

separability would require one coordinate system q_1, q_2 for the first pass $J_{(1)}^*$, another Q_1, Q_2 for the second $J_{(2)}^*$. Once there are two annuli $\mathfrak{U}_1, \mathfrak{U}_2$ forming a torus \mathfrak{T}^2 , everything gets split: \mathbf{q} becomes two points

$$\begin{aligned}\mathbf{q}_1 &\in \mathfrak{U}_1 \subset \mathfrak{U}_1 \\ \mathbf{q}_2 &\in \mathfrak{U}_2 \subset \mathfrak{U}_2,\end{aligned}$$

with

$$\begin{aligned}\mathbf{p}(\mathbf{q}_1, t_1) = \mathbf{p}_1(\mathbf{q}_1) &= dJ(\mathbf{q}_1, t_1) = dJ_{(1)}^*(\mathbf{q}_1) \\ &= \frac{\partial J_{(1)}^*}{\partial q_1} dq_1(\mathbf{q}_1) + \frac{\partial J_{(1)}^*}{\partial q_2} dq_2(\mathbf{q}_1)\end{aligned}$$

at \mathbf{q}_1 and

$$\begin{aligned}\mathbf{p}(\mathbf{q}_2, t_2) = \mathbf{p}_2(\mathbf{q}_2) &= dJ(\mathbf{q}_2, t_2) = dJ_{(2)}^*(\mathbf{q}_2) \\ &= \frac{\partial J_{(2)}^*}{\partial Q_1} dQ_1(\mathbf{q}_2) + \frac{\partial J_{(2)}^*}{\partial Q_2} dQ_2(\mathbf{q}_2)\end{aligned}$$

at \mathbf{q}_2 . Separability is restored by adapting the coordinates

$$\begin{aligned}q_1, q_2 &: \mathfrak{U}_1 \rightarrow \mathbb{R} \\ Q_1, Q_2 &: \mathfrak{U}_2 \rightarrow \mathbb{R}\end{aligned}$$

to J^* , thus satisfying τ . An example would be alignment of q_1 along the level curves of $J_{(1)}^*$, with q_2 perpendicular and appropriately normalised, so that $\partial J_{(1)}^*/\partial q_1$ vanishes and

$$\mathbf{p}_1 = dJ_{(1)}^* = dq_2;$$

and likewise on the second pass:

$$\mathbf{p}_2 = dJ_{(2)}^* = dQ_2.$$

The *double* condition

$$\frac{\partial J_{(1)}^*}{\partial q_1} = 0 = \frac{\partial J_{(2)}^*}{\partial q_1}$$

would be very restrictive and hard to satisfy.

7 Final remarks

Einstein situates his mechanics pretty firmly in the q_i -*Raum* he so often refers to; symplectic abstractions seem rather foreign to him. But one wonders how to make sense of the *configuration* space integrability referred to in the two cases at the bottom of

p. 87. Integrability is closely related to conservation, which surely depends on position *and momentum*. Einstein appears to have the right modern geometrical intuitions—enclosing motion in intersections of level sets of appropriate functions—but in the wrong space. His symplectic ignorance makes *Quantensatz* all the more impressive; a real feat of economy, efficiency: making relatively little go a very long way.

The ‘complete’ integrability one comes across in the modern literature is hardly complete by Einstein’s standards, being compatible with both ergodic and periodic motions. The more extreme kind of integrability Einstein has in mind corresponds to closed, not ergodic motions: *tertium non datur*. He needs a self-intersecting motion assigning more than one momentum here and there—to justify the enlarged configuration space whose topological peculiarities are captured by the homotopy classes he integrates over—but not the *infinitely many* momenta of ergodic motion; his *Riemannisierung* is necessarily finite. And once on the torus, where the greater invariance of Einstein’s integrals $\langle \mathbf{p}, \mathbb{H}_i \rangle$ is evident, the score against Sommerfeld is (an admittedly mathematical, nonempirical) 2-1.

I thank Ermenegildo Caccese, Sandro Graffi, Stefano Marmi and Nic Teh for many useful conversations; and audiences in Urbino, at Paris Diderot and Notre Dame for valuable feedback.

References

- Appell, P. (1909) *Traité de mécanique rationnelle (tome premier)*, Gauthier-Villars, Paris
- Appell, P. (1958) *Traité de mécanique rationnelle (tome deuxième)*, Gauthier-Villars, Paris
- Arnol’d, V. (1988) *Mathematische Methoden der klassischen Mechanik*, Springer, Basel
- Benci, V. & D. Fortunato (2014) *Variational methods in nonlinear field equations*, Springer, Heidelberg
- Born, M. (1925) *Vorlesungen über Atommechanik*, Springer, Berlin
- Born, M. (1933) *Moderne Physik*, Springer, Berlin
- Brillouin, L. (1938) *Les tenseurs en mécanique et en élasticité*, Masson, Paris
- Brogie, L. de (1924) *Recherches sur la théorie des quanta*, Thèse, Paris
- Calin, O. & D. Chang (2005) *Geometric mechanics on Riemannian manifolds*, Birkhäuser, Boston
- Carathéodory, C. (1937) *Geometrische Optik*, Springer, Berlin
- Einstein, A. (1917a) “Zum Quantensatz von Sommerfeld und Epstein” *Deutsche physikalische Gesellschaft, Verhandlungen* **19**, 82-92

- Einstein, A. (1917b) letter to Paul Ehrenfest dated 3 June 1917
- Einstein, A. (1917c) “Eine Ableitung des Theorems von Jacobi” *Preußische Akademie der Wissenschaften, Sitzung der physisch-mathematischen Klasse vom 22. November 1917*, 606-8
- Epstein, P. (1916a) “Zur Quantentheorie” *Annalen der Physik* **50**, 168-88
- Epstein, P. (1916b) “Zur Theorie des Starkeffektes” *Annalen der Physik* **50**, 489-520
- Fasano, A. & S. Marmi (2002) *Meccanica analitica*, Boringhieri, Torino
- Graffi, S. (2004) *Alcuni aspetti matematici della meccanica quantistica*, Quaderni dell’Istituto nazionale di alta matematica, Gruppo nazionale di fisica matematica
- Graffi, S. (2005) “Le radici della quantizzazione” *Quaderni di fisica teorica, Università degli studi di Pavia, Dipartimento di fisica nucleare e teorica*
- Gutzwiller, M. (1990) *Chaos in classical and quantum mechanics*, Springer, New York
- Hamilton, W. (1833) “On a general method of expressing the paths of light, and of the planets, by the coefficients of a characteristic function” *Dublin university review and quarterly magazine* **1**, 795-826
- Hamilton, W. (1834) “On the application to dynamics of a general mathematical method previously applied to optics” *Report of the fourth meeting of the British association for the advancement of science; held at Edinburgh in 1834*, 513-8
- Huygens, C. (1690) *Traité de la lumière*, Vander, Leyden
- Jacobi, C. (1884) *Vorlesungen über Dynamik*, Reimer, Berlin
- Jacobi, C. (1996) *Vorlesungen über analytische Mechanik*, Vieweg, Braunschweig
- Keller, J. (1958) “Corrected Bohr-Sommerfeld quantum conditions for separable systems” *Annals of physics* **4**, 180-8
- Lanczos, C. (1952) *The variational principles of mechanics*, University of Toronto press
- Landau, L. & E. Lifschitz (1970) *Mechanik*, Vieweg, Braunschweig
- Levi-Civita, T. & U. Amaldi (1974) *Lezioni di meccanica razionale (volume secondo, parte seconda)*, Zanichelli, Bologna
- Lowenstein, H. (2012) *Essentials of Hamiltonian dynamics*, Cambridge University Press
- Pettini, M. (2007) *Geometry and topology in Hamiltonian dynamics and statistical mechanics*, Springer, New York
- Rham, G. de (1931) *Sur l’analyse situs des variétés à n dimensions*, Thèse, Paris

- Schrödinger, E. (1926) "Quantisierung als Eigenwertproblem: zweite Mitteilung" *Annalen der Physik* **79**, 490-526
- Sommerfeld, A. (1915) "Zur Theorie der Balmerischen Serie" *Sitzungsberichte der mathematisch-physikalischen Klasse der K. B. Akademie der Wissenschaften zu München* **1**, 425-58
- Sommerfeld, A. (1916) "Zur Quantentheorie der Spektrallinien" *Annalen der Physik* **51**, 1-94
- Spivak, M. (2010) *Mechanics I*, Publish or perish
- Sternberg, S. & V. Guillemin (1984) *Symplectic techniques in physics*, Cambridge University Press
- Weyl, H. (1939) *The classical groups*, Princeton University Press
- Whittaker, E. (1917) *A treatise on the analytical dynamics of particles and rigid bodies*, Cambridge University Press