On Gauge Symmetries, Indiscernibilities, and Groupoid-Theoretical Identities

Gabriel Catren Laboratoire SPHERE - *Sciences, Philosophie, Histoire* (UMR 7219) Université Paris Diderot - CNRS, Paris.

WE-Heraus Seminar on « Hundred Years of Gauge Theory »

Physikzentrum Bad Honnef, July 31st, 2018

We shall critically address a widespread stance

in contemporary (philosophy of) physics,

namely the thesis according to which gauge symmetries

are deprived of any ontological or physical content.

Gauge symmetries would be nothing but a mere mechanism

by means of which we can get rid

of the « *surplus structure* » (Redhead) or the *mathematical/descriptive redundancy*

associated to the different re-coordinatizations

of a unique physical (i.e. coordinate-independent) situation

described by gauge invariant quantities

(Dirac observables, i.e. functions on the reduced phase space).

According to this *« transcendental » position*

gauge symmetries should not be attributed to « nature in itself »,

but rather to the *symbolic means*

by means of which

we constitute an objective description of nature.

« Transcendental » approaches

to the problem of understanding gauge symmetries

might be a *philosophical obstacle*

to the comprehension of the gauge revolution,

even if they did play an important role in the early days of gauge theories

in the framework of Weyl's « purely infinitesimal program »

(c.f. T. Ryckman, *The Reign of Relativity* & his talk).

The « transcendental » interpretation of gauge symmetries

has to face the problem of explaining

the physical consequences of gauge symmetries:

Heuristic gauge argument (local gauge invariance dictates the introduction of gauge fields, albeit not still dynamical).

.Topological solutions (e.g. monopoles, instantons, etc.).

.Relation between gauge symmetries and *renormalization*.

.An uncareful removal of gauge symmetries might lead to pathological or *« bad » quotient spaces*.

Conjecture:

This resistance to accept

that gauge symmetries might have

a physical/ontological content is a particular instance

of a more general « epistemological obstacle »

in the history of philosophy

associated to a moment in which philosophers were confronted

to a particular form of « irrationality ».

Leibniz's Question

Can two numerically different entities be

qualitatively identical?

Leibniz's Principle of Identity of Indiscernibles (PII)

states that such differences « *solo numero* »

are not possible:

Every numerical difference has to result from a qualitative difference.

Perfect copies or clones of a given entity are forbidden.

Concomitant question: *What does confer individuality upon entities?*

According to PII,

intrinsic properties or qualities provide a sufficient criterion to individuate entities,

i.e. there is no further individuation beyond the individuation provided by qualities.

Now, as Leibniz noted,

this principle is not endowed

with any form of a priori logical necessity

(we can conceive possible worlds

with entities differing solo

numero; c.f. M. Black example)

Moreover, Leibniz himself

refused the validity of PII

for mathematical entities (!!!):

``[...] *in nature, there cannot be two individual things that differ in number alone*. And thus, perfect similarity is found only in incomplete and abstract notions, where things are considered [*in rations veniunt*] only in a certain respect, but not in every way, as, for example, when we consider shapes alone, and neglect the matter that has shape. And so it is justifiable to consider two similar triangles in geometry, even though two perfectly similar material triangles are nowhere found."

However, there seems to be a strong resistance in philosophy of physics and mathematics to accept the possibility of rejecting PII. Now, Leibniz's question is intimately related

to a central problem in (the philosophy of) mathematics,

namely the problem of understanding propositions of the form

$$a = b$$
,

i.e. propositions that state

— somehow paradoxically —

that two different things are equal.

And most of the interpretations of statements of the form a = bpreserve the validity of PII:

-nominalistic (Frege, Quine):

a and b are different names of the same thing.

-transcendental (Kant):

a and b denote different ways of thinking about the same thing.

-equivalence relations (Whitehead, Bradley):

a and b are qualitatively different things

that are equivalent only in a certain respect.

None of these interpretations consider expressions

of the form a = b at face value:

stating that two (numerically) different things

are equal.

Moreover, the very set-theoretic « foundations » of mathematics

does not allow for propositions of the form

$$\mathbf{A} = \mathbf{B}$$

for two numerically different sets.

Indeed, if A and B denote sets, either

 $A \neq B$

or

A=B

(if they have the same elements),

which means that there are not in reality two sets,

but only one which may be referred to as either A or B.

All in all, *endorsing PII*

entails an epistemic-transcendental trivialization

of propositions of the form

a = b

But, why does it seem to be so important

to cling to this principle?

Why some philosophers are so resistant

to accept the possibility

of differences « solo numero »?

If PII were not valid,

then we would be confronted

to indexical differences

that are not grounded

on qualitative or property-based differences.

In Wittgenstein's terms,

we would be confronted to *differences*

that can be shown or exhibited

but that cannot be said.

We would be confronted

to the very limits of language !

In Adam's and Weyl's terms,

we would be confronted to an irreducible *thisness* (exposed by means of an *« intuitive exhibition »*)

that cannot be translated into a *suchness*

(defined by a « conceptual determination »).

Rather than taking this *irreducible thisness* at face value

and fully accepting this ontological breakdown of PII,

Weyl understands this irreducible « thisness »

as a « residue » of the phenomenological ego,

i.e. of the subject of *absolute immediate experience* out of which

objective knowledge is constituted

by means of « *symbolic construction* ».

« A conceptual fixation of points by labels [...] that would enable one to reconstruct any point when it has been lost, is here possible only in relation to a coordinate system, or frame of reference, that has to be exhibited by an individual demonstrative act. The objectification, by elimination of the ego and its immediate life of intuition, does not fully succeed, and the coordinate system remains as the necessary residue of the ego extinction. »

H. Weyl, Philosophy of Mathematics and Natural Science, p.75.

Weyl's analysis resumes Kant's take on Leibniz's principle: the possibility of having numerically different and qualitatively identical entities (the effective breakdown of PII) results from the transcendental conditions of possibility of human sensibility (space and time), which are heterogenous with respect to the purely conceptual determinations of the human understanding.

« [...] if the object is appearance, then comparison of concepts does not matter at all; rather, however much everything regarding these concepts may be the same, yet the difference of the locations of these appearances at the same time is a sufficient basis for the *numerical difference* of the object (of the senses) itself. [...] Leibniz took appearances to be things in themselves, and hence to be *intelligibilia*, i.e., objects of pure understanding [...]; and thus his principle of the *indistinguishable* (principium identitatis indiscernibilium) could indeed not be disputed. But since appearances are objects of sensibility [...] space itself-as the condition of outer appearances-already indicates plurality and numerical difference. »

I. Kant, Critique of Pure Reason, A264-B320.

Both in Kant and Weyl, the possibility of having

numerical multiplicities

of qualitatively identical objects

is pulled back to the subjective conditions of human experience,

thereby depriving such a possibility of any intrinsic scope.

Rejecting PII seems to entail a commitment to some form of « irrational » *meta-physical* and *pre-linguistic*

principle of individuation

(haecceitas or primitive thisness).

« By no means I would approve the ways of thinking of a man who did not hesitate long to admit my hecceities. »

C.S. Peirce

Now, with the introduction of the notions

of group and groupoid

mathematics moved forward

in the direction of rejecting PII

and making propositions of the form

$$a = a$$
 & $a = b$

non-trivial.

Identity qua identification

So, let's adopt a *constructive stance* by understanding a proposition

not as a *mere truth value* (« ideal » mathematics)

but rather as the *type of its tokens-proofs*,

i.e. as an abstract statement

whose concrete realizations are given by its different proofs.

For instance, what does it mean to say

that two figures in the Euclidean space

are equal?

« If indeed one tries to clarify the notion of equality, which is introduced right at the beginning of Geometry, one is led to say that two figures are equal when one can go from one to the other by a specific geometric operation, called a *motion*. »

E. Cartan.

Every « motion » that transports

one figure *a* on the top of the other figure *b*

is a concrete proof of the fact that the proposition

$$a = b$$

is true.

The fundamental step is to state that two things are equal if they can be concretely *identified* by means of an *isomorphism* (or reversible transformation)

Consequences:

- 1. Against PII, numerically different things might be qualitatively « identical ».
- 2. Two things might be identical *in many different ways*, i.e. propositions of the form a = b might admit many proofs.
- 3. This set of proofs might carry a non-trivial structure (e.g. homotopy theory).
- 4. Far from entailing a slip into irrationality, the very structure of these indiscernibilities can be mathematically formalized.

If we make abstraction of the concrete proofs

of a proposition of the form a = b

(i.e. the identifications between a and b)

by just retaining its truth value

we might lose the non-trivial higher structure

of identifications, identifications between identifications, etc.

Homotopical Maxim:

Never forget that *abstract equalities*

a = b

are underpinned by concrete identifications

$$a \xrightarrow{\simeq} b$$

Galoisian Principle

The first step in this direction was given by Galois in the beginning of the XIX century, when he tackled a similar problem namely the existence of solutions to polynomial equations formulated in a field K such that no K-relation can discern them.

Rather than trying to « save » rationality

by forbidding indiscernibilities, Galois

(1) fully accepted that we can have numerical differences

not grounded on qualitative differences,

(2) and introduced a mathematical notion (groups)

that formalizes the structure of that which cannot be said,

namely the numerical difference between indiscernible solutions.

Remark

In the framework of Galois theory

these indiscernibilities are purely epistemic rather than intrinsic

since they can be broken by passing to a larger field

endowed with a higher « resolving power »:

What cannot be said in a given language

becomes sayable in an extension of that language.

But more generally,

any mathematical structure

endowed with non-trivial automorphisms

is intrinsically composed

of numerically different

and indiscernible elements

For instance, the numerically different points

of the Euclidean plane

carry no intrinsic qualities

by means of which we could discern or individuate them.

The Plasticity of the Identity

Now, by following the prescription according to which

we have to substitute the notion of equality

by the notion of identification qua isomorphism,

each automorphism of an entity a

can be understood as a non-trivial proof of the fact that

the *identity principle* a = a is true.

a is identical with a if there exists transformations that reversibly transform a into a.

Since there is always the trivial transformation, the proposition a = a is always true.

.But if the structure has automorphisms, then there are non-trivial transformations \ll moving $\gg a$ into itself.

The identity principle *a* = *a* understood as the type of its tokens-proofs carries information about the symmetries of *a*, thereby becoming an entity-dependent proposition.

And the notion of *group* formalizes

this « factorization » of the trivial identity of a non-rigid entity a

in many non-trivial self-identifications

given by endomorphisms $f: a \to a$

satisfying

$$id_a = f \circ f^{-1}$$

The « Relativity Problem » According to Weyl

The fact that automorphisms

have to be understood as *intrinsic* features of the entity at stake

was stressed in different forms by Weyl.

According to Weyl (PMNS),

we have to distinguish two phases of (what he calls)

the « relativity problem ».

First, there is an ontological dimension

given by the fact that a structure

might be *intrinsically* endowed

with non-trivial automorphisms,

which means (in our terms) that its identity might be non-trivial.

Second, there is an epistemic dimension

of the « relativity problem »

given by the fact that these automorphisms

« objectively distinguish » a class of coordinate systems.

It follows from Weyl's analysis

that the requirement of invariance under frame changes

is a derived "epistemic" consequence

of the *intrinsic* symmetries of the structure.

For instance, the problem of (special) relativity,

rather than being a « transcendental » problem

concerning the subjective conditions of physical objectivity,

« [...] is nothing else than the problem of dealing with the *inherent* symmetry [encoded by the Poincaré group] of the four-dimensional continuum of [flat] space and time. »

H. Weyl, *Symmetry*

« The physicist will question Nature

to reveal him her true group of automorphisms »

H. Weyl (quoted by E. Scholz)

« Whenever you have to do with a structure-endowed entity S try to determine its group of automorphisms, the group of those elementwise transformations which leave all structural relations undisturbed. You can expect to gain a deep insight into the constitution of S in this way."

H. Weyl, Symmetry.

Riemann-Cartan Principle

Now, symmetric structures are very particular kinds of structures

(e.g. symmetric solutions to the Einstein's equations).

So, why do they play such a central role

in the gauge theories of fundamental interactions?

The answer was announced by Plato

(the « elements » are given by regular structures; c.f. Timaeus),

prepared by Riemann,

and acquired a full-fledged development thanks to E. Cartan.

Symmetric structures acquire a universal scope

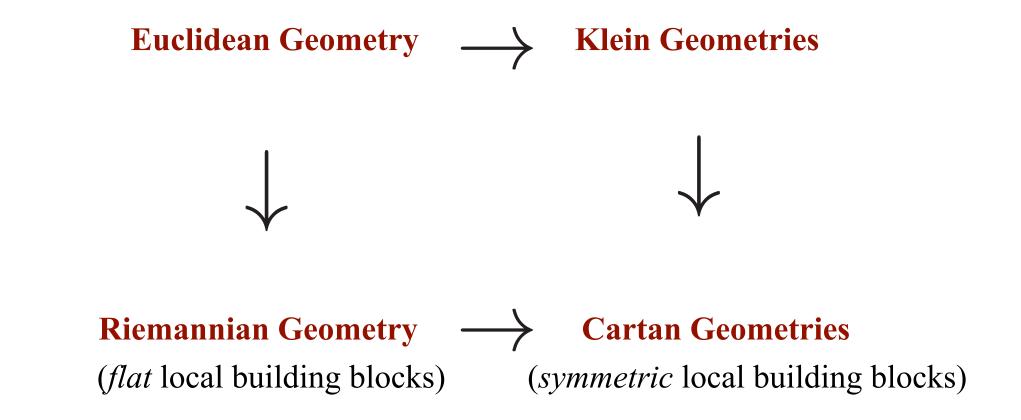
given by the fact that they can be understood as

local building blocks

that can be « connected » in a « curved » manner

in order to produce more general structures,

i.e. structures that are not necessarily symmetric.



C.f. Sharpe, R.W. (1997). *Differential geometry. Cartan's generalization of Klein's Erlangen program*. NY: Springer-Verlag.

Cartan's twofold generalization of both Riemannian geometry and Klein's Erlangen program goes against Weyl's (and Cassirer's) attempt to save Kant's transcendental aesthetics from the menace posited by GR.

Briefly, Weyl claimed that only the *infinitesimal spatiotemporal neighborhood* of the experiencing subject has to carry an *a priori* geometric structure (Euclidean, Minkoswki).

However, Cartan's approach

— by stressing the Kleinian notion of *symmetry*

to the detriment of the Riemannian notion of *flatness* —

allows for the possibility of having

different infinitesimal symmetric models.

According to the *Riemann-Cartan principle*,

the construction of general structures

requires to have copies of a symmetric structure.

Since by definition of the notion of « copy »,

copies are

numerically different and qualitatively identical,

the Riemann-Cartan principle

requires to suspend the validity of PII.

In order to formally deal

with families of identical elements,

the notion of group

(encoding the multiple self-identifications of a single entity)

is not enough.

A classical mathematical notion

that does part of the job

is the notion of *equivalence relation*

(a *reflexive*, *symmetric*, and *transitive* relation)

in a set of entities.

However, the limit of equivalence relations results from the fact

that two entities *x* and *y* are either equivalent

 $x \sim y$

or not,

but they cannot be equivalent in many different ways.

In particular

 $x \sim x$

in just one way.

However, if x and y are two copies of a non-rigid entity, then

.both x and y are identical to themselves in different ways,

.there are many identifications between *x* and *y*.

In a sense, groups and equivalence relations are « orthogonal » notions:

Whereas groups formalize

single objects with many self-identifications,

equivalence relations formalize

sets of many objects endowed with unique identifications.

So, we need a mathematical notion capable of encoding

.multiple identifications

.between *multiple objects*.

Briefly, a groupoid

 $s, t: G_i \rightrightarrows G_o$

is a family of objects G_o

endowed with a family of identifications G_i

(*reversible* transformations)

between some of them such that

1) each object a is endowed with a trivial identity id_a

2) the compositions of composable transformations exist and is associative.

Given a groupoid, we can define the *underlying equivalence relation*, where

 $a \sim b$

if there exists at least one arrow

 $a \rightarrow b$

This equivalence relation makes abstraction from the fact that

a and b might be identified in many different concrete ways

in order to only keep the abstract fact that a and b are identical.

In type-theoretic terms, the expression

 $a \sim b$

can be understood as an abstract proposition (a type)

whose tokens-proofs are given

by all the concrete identifications

$$a \xrightarrow{\simeq} b$$

Yang-Mills Theories

The geometric setting of Yang-Mills theories

is given by bundles

— parameterized by spacetime M —

of identical copies of a non-rigid G-structure S.

A G-bundle is a triple

 $(E \xrightarrow{\pi} M, S, G)$

such that the standard fiber S

is endowed with a G-action

and the transition functions

on the intersections of an open covering $\{U_i\}$ of M

have values in G.

Since the fibers are identical copies

of the non-rigid structure S

we can paste the fibers in non-trivial manners

in the intersections $U_i \cap U_j$

thereby introducing twists in the fibration

(which are the geometric source of the topological solutions)

We can now pass to the second phase

of the « relativity problem »

and introduce the corresponding

G-principal bundles of frames

Each fiber S_x can be identified with the standard fiber S

by means of different non-canonical isomorphisms.

Each isomorphism

 $p: S \to S_x$

defines a particular framing of S_x

To « frame » a fiber S_x

means to choose a concrete identification

between S_x and the standard fiber S.

The action of G on the standard fiber S

induces a right (free and transitive) action

on the set of frames $Iso(S, S_x)$ on each x in M

given by composition

$$(p: S \to S_x) \mapsto (pg: S \xrightarrow{g} S \xrightarrow{p} S_x)$$

which means that each $Iso(S, S_x)$ is a *G*-torsor.

Changes of frames are just

the « epistemic » effect on the frames

 $p: S \to S_x$

of the intrinsic automorphisms

of the standard fiber S.

The collection of all this G-torsors

 $Iso(S, S_x)$

for each x in M defines the

G-principal bundle of local frames

 $P \to M$

The "epistemic" invariance of a gauge theory

under local gauge transformations

is just a consequence of the *intrinsic* fact

that we are dealing with families of identical copies S_x

of a non-rigid structure S.

Connections

From a physical standpoint, if we understand the fibers S_x

as « internal spaces » of states for material fields,

then the fact that S_x are copies of a symmetric entity S

implies that we cannot naturally compare internal states

in different spatiotemporal locations

(since there are no canonical isomorphisms between any two copies of *S*).

One possible solution to this problem

would be to introduce a gauge fixing section

 $s: M \to P$

(an election of an « origin » in each G-torsor $Iso(S, S_x)$)

modulo global gauge transformations.

We could even try to understand such a section as a dynamical field,

thereby avoiding arbitrary elements in the theory.

This solution amounts to « break »

the symmetries of the different copies S_x

in such a way that they become canonically isomorphic.

By doing so, we are no longer under the aegis

of the Riemann-Cartan principle,

which requires the elementary building blocks

to be symmetric structures.

Now, is it possible to endow the bundle

 $E \to M$

with a « transversal coherence »

by means of which we could compare internal states

in different spatiotemporal locations

without « breaking » the intrinsic symmetries

of the identical objects S_x

that is by preserving local gauge invariance?

The answer is yes.

And to do so requires (in our own terms)

to endow the bundle

 $E \to M$

with a *concrete system of identifications*

between the different fibers S_x

Now, we have two different regimes of identifications in the bundle

 $E \to M$

 The identifications between the different points of M encoded in the *path groupoid* of M,
the groupoid of (thin homotopy classes of) paths in M

(which does not take into account the fibers over each point)

2. The identifications between the different identical fibers S_x (which does not take into account the path-dependent identifications between points in M)

In order to unify these two systems of identifications,

it is useful to think about the points x of M

as « fat » or structured points endowed with an internal structure

given by the fibers S_x .

An Ehresmann connection is an enrichment

of the external (spatiotemporal)

concrete path-dependent identifications

between the « thin » or structureless points of M

(encoded in the path groupoid of M)

into identifications

between the « fat » or structured points.

Ehresmann's Groupoid-Theoretical Description

Ehresmann connections can be entirely defined

in groupoid-theoretical terms

that stress the underlying « homotopical »

detrivialization of equalities

made possible by the rejection of PII.

Given a G-principal bundle

 $P \to M$

we can introduce Ehresmann's structural groupoid

 $P \times_G P \rightrightarrows M$

Roughly speaking,

.the objects are the fibers S_x

.the arrows are identifications between any two fibers.

An arrow of the structural groupoid

is given by an equivalence class $[(p_2,p_1)]$

under the action

 $(p_2, p_1) = (p_2g, p_1g)$

Each arrow defines the following identification

between the copies S_1 and S_2 of S: $p_2 \circ p_1^{-1} : S_1 \xrightarrow{p_1^{-1}} S \xrightarrow{p_2} S_2,$ The structural groupoid can be projected

to the underlying maximal equivalence relation

between the points of M

given by the *pair groupoid* of M

 $M\times M\rightrightarrows M$

This groupoid encodes the abstract identities

 $x_1 = x_2$

between any two « thin » points of M

The kernel of the projection

 $P \times_G P \to M \times M$

is given by the so-called gauge groupoid

$$P \times_G G \rightrightarrows M$$

where the right action of G on $P \times G$ is given by

$$(p,g) \cdot h = (ph, h^{-1}gh)$$

We then have the following exact sequence of Lie groupoids

 $P \times_G G \hookrightarrow P \times_G P \twoheadrightarrow M \times M$

where the inclusion on the left is given by.

$$[(p,g)]\mapsto [(p,pg)]$$

which means that an arrow

 $[(p,g)] \in P \times_G G$

can be interpreted as a map

$$p \circ (pg)^{-1} : S_x \xrightarrow{p^{-1}} S \xrightarrow{g^{-1}} S \xrightarrow{p} S_x$$

i.e. as an automorphism of the G-torsor S_{x}

The gauge group of vertical automorphisms of

 $P \to M$

can be obtained as the group of bisections

of the gauge groupoid,

 $P \times_G G \rightrightarrows M$

where a bisection is (roughly speaking)

a smooth election of a single outgoing arrow for each object x in M,

i.e. of an automorphism of S_x for each x.

Now, the projection $P \times_G P \to M \times M$ relates the concrete identifications between the fibers S_x

to the abstract identities between points of M,

and not (as we wanted) with the concrete identifications

between points of M

given by (classes of) paths in M.

This problem can be solved

by passing to an infinitesimal description

and integrating along paths in M.

The sequence of Lie groupoids

 $P \times_G G \hookrightarrow P \times_G P \twoheadrightarrow M \times M$

gives rise

— by passing to an infinitesimal description —

to the following exact sequence of Lie algebroids

 $P \times_G \mathfrak{g} \to TP/G \xrightarrow{\pi} TM$

An *Ehresmann connection* is a section

$$\gamma:TM\to TP/G$$

of the map

 $TP/G \xrightarrow{\pi} TM$

The connection γ « connects »

the different fibers by selecting,

for each infinitesimal abstract identification v in $T_x M$

an infinitesimal concrete identification $\gamma(v)$

between the fibers E_x and E_{x+v}

Conclusions

The notions of group and groupoid convey

a rejection of PII

and, consequently,

a « de-trivialization » of statements of the form

a = a and a = b

respectively.

The concrete self- and hetero-identifications in a family of entities

might carry a complex (homotopical) structure.

By flattening these concrete identifications

into abstract (or mere) propositions of the form

$$a = a$$
 and $a = b$

we loose track of all this information.

Gauge theories provide a particular realization of these ideas

in the form of the *Riemann-Cartan principle*,

which states that global general structures can be constructed

by « connecting » numerically different and qualitatively identical

local symmetric building blocks.

Local gauge symmetries

— far from resulting from a « transcendental » constraint on scientific knowledge

(objectivity qua coordinate invariance) —

are a natural consequence of the fact

that the local building blocks

are numerically different and identical copies

of a non-rigid structure.

If we consider gauge symmetries as mere mathematical redundancies

that can be simply removed

by taking the quotient by the action of the gauge group,

then we are led to pathological or « bad » quotient spaces.

The most powerful method to deal with gauge symmetries (the BRST-cohomological reformulation

of the *restriction* + *quotient operation*

by means of which one defines the reduced phase space of a constrained Hamiltonian theory)

is an equivariant method

that does not only takes into account the identifications between

different « representatives » of the gauge fields,

but also the identifications between identifications (« ghosts of ghosts »)

and so on and so forth all the way up.

Contemporary (homotopical) mathematics, like

.higher category theory,

.homotopy type theory,

.derived geometry.

is moving forward

(by defining in particular « good » quotient spaces)

by unpacking and exploring

the structure of (higher) identifications

enveloped by statements of the form a = b.

By paraphrasing B. Timmermans

(*Histoire philosophique de l'algèbre moderne*)

we could say that it might be utterly misleading to "conceal the *pluralism of* [gauge] *transformations* [by considering them as mere mathematical redundancies associated to the symbolic constitution of objective knowledge] for the sole benefit of the *realism of invariants*"

Many thanks !

References:

-G.C, Klein-Weyl's Program and the Ontology of Gauge and Quantum Systems, Studies in History and Philosophy of Modern Physics, 2017.

-G.C., Geometric Foundations of Cartan Gauge Gravity, International Journal of Geometric Methods in Modern Physics, 2015.

-G.C & J. Page, On the Notions of Indiscernibility and Indeterminacy in the Light of the Galois–Grothendieck Theory, Synthese, 2014.

-G.C, On the Relation Between Gauge and Phase Symmetries, Foundations of Physics, 2014.

-G.C, Geometric Foundations of Yang-Mills Theories, Studies in History and Philosophy of Modern Physics, 2008.

-G.C. & J. Devoto, Extended Connection in Yang-Mills Theory, Communication in Mathematical Physics, 2008.