Derivation of Lorentz transformation from principles of statistical information theory

Revised 2/28/2016

Thomas E. Butler
1 Woodruff Way, Columbia, New Jersey 07832
e-mail address: astraclara@embarqmail.com

The Lorentz transformation is derived from invariance of an information quantity related to statistical hypothesis testing on single particle system identification parameters. Invariance results from recognition of an equivalent observer as one who reaches the same conclusions as another when the same statistical methods are used. System identity is maintained by parameter values which minimize discrimination information, given by a Kullback-Liebler divergence, under a constraint of known shift in observation time. Deviation of discrimination information from the minimum value gives the difference in information between an observed system under a constraint shift and the expected system that maintains identity under the same constraint. System observation states are represented by parametric probability distributions of particle system measurement values.

Keywords: Lorentz transformation, special relativity, information, Kullback-Liebler discrimination information
PACS: 03.30.+p;89.70.Cf;01.55.+b

I. INTRODUCTION

The observation state of a free single particle system at a constant velocity is represented by a probability distribution of observation measurement values. Change in the observation time results in a change of the observed system state to a state that retains the system identity of the initial system. The system identity [1] [2] [3] is preserved by a final observation state that minimizes the number of bits of discrimination information given by the Kullback-Liebler discrimination information [4] under the constraint of an observation time shift. Excess discrimination information above the minimization value shows how close an observed system state is to a state which preserves system identity, and is an input to identity hypothesis testing methods. Different observers are expected to obtain the same system identity hypothesis conclusions when using the same statistical methods. Then the state associated with a different observer must keep the excess discrimination invariant. Invariance of the excess discrimination information for a free single particle system gives the Lorentz transformation, not of random variables which do not transform, but of the observed system state as represented by parameters of the probability distribution.

Derivations of the Lorentz transformation from within other areas of physics, for example quantum information and communication theory [5], make the transformation dependent on theories within those branches of physics. The derivation based on a discrimination invariant shares a similar dependence, but primarily on concepts of statistical information theory, independent of other branches of physics beyond representation of simple Galilean motion. Although the statistical discrimination derivation depends on specific probability models, the resultant transformation is the same for all similar probability models for small shifts in space and time parameters. The information based derivation also supports generalization of the Lorentz transformation to a concept of equivalence transformations of distributions of arbitrary quantities.

Another characteristic of the statistical information derivation of the Lorentz transformation is that there is no assumption of an invariant speed [6] [7]. Domain and range set properties of equivalence transformations imply existence of an upper bound on the magnitude of the velocity. Then transformations that preserve excess discrimination invariance require all boundary velocities to map to the boundary.

II. PARTICLE OBSERVATION STATE

Observations of the motion of a particle in an experiment produce a set of times of observation and associated particle positions. In a repeatable experiment, the information contained by an ensemble of sets of position-time value pairs can be concisely represented by a probability distribution which normalizes over both observation time and position random variables, as shown in the random sample of Figure 1. Each experiment is independent, so the probability distribution represents the expected distribution of a single experiment. No special

interpretations are required beyond the probability distribution as a simple representation of measurement data, given that a measurement exists. The probability distribution is the observation state of the particle system.

An observation time and position density that represents measurement data is applicable to many variants of a position versus time experiment. A single experiment can consist of dynamic observations in which the particle time and position are recorded as the particle moves. An alternative experiment might consist of a single time-position measurement, with apparent motion from an ensemble of repeated experiments with different observation times. The same density represents measurements for both interpretations.

A. Example probability density

The observed particle state is a probability density in position and time random variables which provides a realistic summary of observation data from a controlled experiment. This is made evident by a plot of data generated by a simulated random sample from a normal time and space probability density, shown in Figure 1.

The probability density of Figure 1, which is normalized over both observation time and position random variables, generates a random sample that appears to be an ordinary collection of observation data from a repeated constant velocity experiment of finite duration that might have come from a physics class lab. The same result could also be produced by a single experiment with a large number of observations. In either case the measurement position standard deviation of 16.667 is not an error, but primarily the width of the range of the main concentration of position observations. Position error is given by the conditional position given time standard deviation of 0.745. Similarly, the measurement time standard deviation of 2 is not an error but primarily the extent of the range of observed times. Peak frequency of observations increases without limit for a dynamic interpretation of the model, as the number of observations rises, without any effect on the probability density. In this respect the normal density exhibits a classical physics behavior of independence from arbitrarily small time intervals.

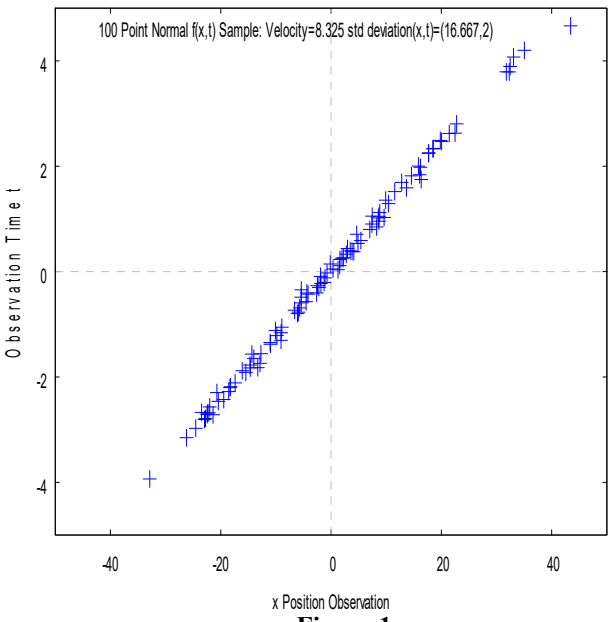

Figure 1

B. A model of uniform motion

The example probability density of Figure 1 is a two dimensional normal probability density in observation position x and time t random variables as a model of experimental data from non-quantum mechanical uniform particle motion. Extension to a normal density in four dimensions models a three dimensional velocity. The four dimensional normal density f is

$$f\left(R:\bar{R},\Sigma\right)=\frac{1}{\sqrt{\det\left(2\pi\Sigma\right)}}e^{-\frac{1}{2}\left(R-\bar{R}\right)^{T}\Sigma^{-1}\left(R-\bar{R}\right)}, \quad (1)$$

with random variable vector $R=\begin{pmatrix}x & y & z & t\end{pmatrix}^{T}$, parameters mean vector $\bar{R}=\begin{pmatrix}\bar{x} & \bar{y} & \bar{z} & \bar{t}\end{pmatrix}^{T}$, and covariance matrix Σ. Define the vector V by $V=\sigma_t^{-2}\begin{pmatrix}\Sigma_{xt} & \Sigma_{yt} & \Sigma_{zt}\end{pmatrix}^{T}$. Then the position-time block matrix form of Σ is

$$\Sigma=\begin{pmatrix}C+\sigma_t^{2}VV^{T} & \sigma_t^{2}V\\ \sigma_t^{2}V^{T} & \sigma_t^{2}\end{pmatrix}, \quad (2)$$

with C the 3x3 position given time co-variance matrix and σ_t^{2} is the observation time variance. Define the vector of random variable observation position components by

$X = \begin{pmatrix} x & y & z \end{pmatrix}^T$ with mean $\overline{X} = \begin{pmatrix} \overline{x} & \overline{y} & \overline{z} \end{pmatrix}^T$ The density *f* is the product of the marginal probability density g in the observation time random variable and the conditional probability density h in observed position given observed time, which for (1) are defined by

$$h\left(\varepsilon : \mathrm{C}\right) = \frac{1}{\sqrt{\det\left(2\pi C\right)}} e^{-\frac{1}{2}\varepsilon^T C^{-1}\varepsilon}, \text{ and}$$

$$g\left(\delta t : \sigma_t\right) = \frac{1}{\sqrt{2\pi}\sigma_t} e^{-\frac{\delta t^2}{2{\sigma_t}^2}}. \quad (3)$$

with $\varepsilon = \delta X - V\delta t$, $\delta X = X - \overline{X}$, $\delta t = t - \overline{t}$ and where

$$f\left(R : \overline{R}, \Sigma\right) = m\left(\delta t : \sigma_t\right) h\left(\varepsilon : \mathrm{C}\right) \quad (4)$$

. Although the position standard deviation of (4) is the width of the range plus position error, and is not a position error, the diagonal components the position covariance matrix C are an indication of position error. The four dimensional probability density with parameters $\overline{X}, \overline{t}, \mathrm{V}, C$ and σ_t is selected as a model of single free particle motion.

III. PARAMETERS

Parameters of the probability distribution are a representation of the particle observation state. A theory provides the general form of the distribution as a collection of possible allowed distributions, with parameter values to select a particular distribution. All inputs allowed by the applicable theory which select a particular $f\left(R : \overline{R}, \Sigma\right)$ from the set of all possible densities supported by the theory are parameters. Parameters are constructed such that there is a one to one relationship between a parameter value and a probability distribution. Initial, boundary and environmental conditions of an experiment are parameters, as are numerical parameters. Statistical methods are used to estimate parameter values from the output data of an experiment.

Parameters of probability distributions are a necessary and common component of physics. Pressure, temperature, entropy and density are thermodynamic parameters associated with probability distributions of the underlying statistical mechanics of very large degree of freedom composite systems. Some parameters are designed to be controlled by an experiment, and others set free to vary in response to the controlled parameters. Thermodynamics provides the similar example that for the same ideal gas equation of state an isentropic gas produces a different relation between pressure and density than that of an isothermal gas. Parameters can be classified as either constraints, controlled by an experiment, or as unconstrained, left to be determined as parameters of a distribution of measurements after an experiment is performed.

A. Parameter classification

All parameter values are extracted from experimental data. Prior to execution of an experiment the design of an experiment can select planned values of some parameters, based on prior theoretical considerations independent of the outputs of the experiment. Mean and extent of the range of the observation time are two parameters which might be set by the design of an experiment, and confirmed by experimental data, since they can be regarded as under the control of the observer and independent of the object particle of the experiment. Parameters are classified as constraint or responsive in a given experiment:

1. *Constraint Parameters*

Constraint parameters are parameters with planned values input to an experiment which exhibit high levels of repeatability and are independent of the subject of an experiment. The value of a constraint associated with a constraint parameter value is given by a subset of the set of allowed parameters to which the constraint confines distribution selection.

2. *Responsive Parameters*

Responsive parameters have values which are not planned input but are outputs from the execution of an experiment and may depend on the subject of the experiment.

Let $P = \begin{pmatrix} \overline{X} & V & \overline{t} & \sigma_t & C \end{pmatrix}$ represent the parameters of the probability density so that (4) is condensed to

$$f\left(R : P\right) = m\left(\delta t : \sigma_t\right) h\left(\varepsilon : \mathrm{C}\right) \quad (5)$$

Some parameters or transformations of parameters might be selected to be either constraints or responsive, for different experiment designs. Other parameters can have intrinsic properties that effectively classify the parameters as constraint or responsive parameters. Time parameters have such intrinsic properties.

4. *Constraint Sets*

Let $\mathcal{Q} = \{\forall P\}$ be the set of all probability distribution parameters allowed by a theory. An experiment design defines constraints which restrict parameter values to a subset $\mathcal{C} \subset \mathcal{Q}$. The value of a constraint is given by subset $\mathcal{C}$. Denote the set of all valid constraint sets as $\Lambda(\mathcal{Q})$, so $\mathcal{C} \in \Lambda(\mathcal{Q})$.

Define the set of constraint sets $\Upsilon(\mathcal{C})$ relative to a set $\mathcal{C}$ to be the set of all subsets of $\mathcal{C}$ which are also valid constraint sets, so that $\Upsilon(\mathcal{C}) = \Lambda(\mathcal{Q}) \cap \mathrm{P}(\mathcal{C})$ where $\mathrm{P}(\mathcal{C})$ is the set of all subsets of $\mathcal{C}$. All constraints associated with a particular type of parameter constraint are contained in an adjustable constraint relative to constraint $\mathcal{D}$ which is defined to be a set $\mathrm{A}(\mathcal{D})$,

$$\mathrm{A}(\mathcal{D}) = \{\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3, \ldots\}\ , \tag{6}$$

with elements $\mathcal{C}_i \subset \mathcal{D}$ that satisfy

$$\mathcal{C}_i \cap \mathcal{C}_j = \varnothing \tag{7}$$

for distinct elements $\mathcal{C}_i \neq \mathcal{C}_j$, and where the union of all elements satisfies

$$\mathcal{C}_1 \cup \mathcal{C}_2 \cup \mathcal{C}_3 \cup \ldots = \mathcal{D} \tag{8}$$

Properties (7) and (8) of an adjustable constraint imply that every element of $\mathcal{D}$ is contained within one and only one element of $\mathrm{A}(\mathcal{D})$. Then there exists a function $M(P : \mathrm{A}(\mathcal{D}))$ for each $P \in \mathcal{D}$ which returns the constraint set element of $\mathrm{A}(\mathcal{D})$ that contains P. For example, if one of the co-ordinates of P is the mean observation time $\overline{t}$ parameter constraint, then index $i = \overline{t}$ and $M(P : \mathrm{A}(\mathcal{D})) = \mathcal{C}_{\overline{t}}$. If a type of constraint is defined by known values of parameter components denoted by ζ then ζ can be used as the index and $\mathcal{C}(\zeta : \mathrm{A}(\mathcal{D}))$ is the corresponding constraint set element of the adjustable constraint set.

Let $\mathcal{C}_i \in \mathrm{A}_c(\mathcal{D})$ be an element of an adjustable constraint and $\mathcal{B}_i \in \mathrm{A}_b(\mathcal{D})$ be an element of a different adjustable constraint. Then the elements of set $\mathrm{A}_{(b,c)}(\mathcal{D}) = \{\mathcal{B}_i \cap \mathcal{C}_j \ \forall \mathcal{B}_i \neq \mathcal{C}_j\}$ satisfy (7) and (8) so that $\mathrm{A}_{(b,c)}(\mathcal{D})$ is an adjustable constraint relative to $\mathcal{D}$.

B. Confirmation Sets

Verification of a theory is rarely accomplished by a single type of experiment represented by the constraint choices in a one adjustable constraint set. More typical is a variety of types of experiments to provide a stronger verification of a theory. Each type of experiment corresponds to a different adjustable constraint set. Define a confirmation set of adjustable constraints to be the set of all adjustable constraints used in experiments to verify a theory.

Designate $\tilde{A}$ given by

$$\tilde{A} = \{\mathrm{A}_1(\mathcal{D}), \ldots, \mathrm{A}_n(\mathcal{D}), \ldots\} \tag{9}$$

as the confirmation set for the single particle theory.

C. The Time Postulate

In an experiment set up to measure the particle position about a specified time the mean observed time and the extent of the range of observed times give all parameters – mean and standard deviation - of the normal marginal time density. Since all of the time measurements are under control of the observer, all of the marginal observation time density parameters are constraint parameters.

Set up a different experiment in which a clock is triggered to measure the time the particle passes a detector at a known position. This experiment can be used to measure mean velocity, and provides an example where the mean position is a constraint parameter and the mean time is not a constraint.

A verifiable theory with a time and spatial components includes adjustable constraints corresponding to both types of experiments, which are elements of the confirmation set of the theory. A prominent characteristic of observation time distributions is the tendency whenever possible to regard the time distribution parameters as experimental constraints. This tendency is prevalent because time values are generally considered to be under the control of the observer to the maximum extent possible. These consideration suggest the time postulate, which is

1. *Time Postulate*

Every parameter that selects a marginal distribution of an observation time independent of any other random

observation variables is a control parameter, and thus an experimental constraint, in at least one Adjustable Constraint set element of a Confirmation Set of a theory with time dependence, and it is this quality of the necessity of control that distinguishes time from all other quantities.

To the extent that every mechanism is repeatable and controlled, all mechanisms are clocks. Under the time postulate, parameters which describe a clock's data must all be under the control of an observer in at least one verification experiment, and are therefore constraint parameters for that experiment.

Each confirmation set of a time dependent theory must contain at least one adjustable constraint in which all parameters of the marginal observation time density are constraint parameters. The design of the experiment must select one of these time parameter adjustable constraints as a *primary constraint,* to serve as the focus of statistical decisions. Only the primary constraint is used to determine the excess discrimination invariant. This is because only a constrained time parameter marginal density provides potential operator selection of all possible time measurement scenarios that are under control of the observer with certainty. Parameters in other adjustable constraints in a confirmation Set are incomplete in that they do not provide certainty of access to all possible measurement times, for if they did the time postulate requires that they also be time, potentially as a parameter of a clock mechanism.

IV. CHANGE OF OBSERVATION STATE

A. Parameter Shifts

A single particle probability density in observation time and position random variables models data from an experiment of finite duration. Motion can continue outside the range of data from an experiment, which presents the opportunity for observation data collected by additional experiments. The density for each experiment is represented by different parameter values.

Particle observation state is given by the probability density which represents data from an experiment. Different densities represent different experiments, and different particle observation states. Motion of a particle through different experiments occurs with a change of particle observation states. Transitions between observation states of the particle are represented as changes in parameters of the probability density.

B. Discrimination Information

A sequence of experiments which measure position and observation time of a single particle produces a sequence of observation data and a corresponding sequence of observation data probability density parameters. Parameter values which are far from values expected for the uniform motion of a free particle are not representative of the uniform motion of the originally observed free particle. When trajectory information is the only available particle identity information then only parameters along an initial trajectory maintain the identity of the particle system.

Let P_0 be the initial particle system observation state parameter, and P_1 the observation state of a subsequent experiment with new constraint parameter values. The discrimination information available in the density in favor that analysis of data selects P_1 over P_0 is

$$\begin{aligned} I\left(P_1,P_0\right) &= \int_{\forall R} dxdydzdt\ f\left(R:P_1\right)\ln\frac{f\left(R:P_1\right)}{f\left(R:P_0\right)} \\ &= \int_{-\infty}^{\infty} dt\ m\left(\delta t_1:\sigma_{t1}\right)\ln\frac{m\left(\delta t_1:\sigma_{t1}\right)}{m\left(\delta t_0:\sigma_{t0}\right)} + \\ &\int_{\forall R} dxdydzdt\ m\left(\delta t_1:\sigma_{t1}\right)h\left(\varepsilon_1:\mathrm{C}_1\right)\ln\frac{h\left(\varepsilon_1:\mathrm{C}_1\right)}{h\left(\varepsilon_0:\mathrm{C}_0\right)} \end{aligned} \tag{10}$$

which evaluates to

$$\begin{aligned} I_{10} = &\frac{1}{2}\frac{\left(\Delta\bar{t}\right)^2}{{\sigma_{t0}}^2} - \ln\frac{\sigma_{t1}}{\sigma_{t0}} - \frac{1}{2}{\sigma_{t1}}^2\Delta{\sigma_t}^{-2} \\ &-\frac{1}{2}\ln\frac{\left|\det C_1\right|}{\left|\det C_0\right|} - \frac{1}{2}C_1\Delta C^{-1} \\ &+\frac{1}{2}{\sigma_{t1}}^2\Delta V^T{C_0}^{-1}\Delta V + \frac{1}{2}\delta\Pi^T{C_0}^{-1}\delta\Pi \end{aligned} \tag{11}$$

where Δ indicates a component of $P_1 - P_0$ and $\delta\Pi = \Delta\bar{X} - V_0\Delta\bar{t}$.

C. Preservation of identity

An initial observation state parameter P_0 shifts to a parameter P_1 for a single particle system when controlled constraint parameters, such as the mean observation time component of P_1, shift to planned values. Adjustable constraint sets provide structures to constraints which can be used by all observers. A particular adjustable constraint set defines possible constraint values for a particular type of constraint associated with an experiment.

Parameters R_1 which preserve identity of a system with observation state parameter P_0 under a shift in constraint value must minimize discrimination information and so must satisfy

$$I\left(R_1, P_0\right) = \min_{p \in \mathcal{C}\left(\zeta_1 : \mathrm{A}_\zeta\left(\mathcal{Q}\right)\right)} I\left(p, P_0\right) \tag{12}$$

where $\mathcal{C}\left(\zeta_1 : \mathrm{A}_\zeta\left(\mathcal{Q}\right)\right) = M\left(R_1 : \mathrm{A}_\zeta\left(\mathcal{Q}\right)\right)$.

The time postulate requires $\bar{t}_1$ and σ_{t1} be constraint parameters. Position error in pre-quantum theory C_1 is under the control of the observer, so C_1 is also a constraint parameter. With $\zeta_1 = \left(\bar{t}_1 \quad \sigma_{t1} \quad C_1\right)$ as constraint parameter, only parameters $\bar{X}_1$ and V_1 remain to be adjusted in (11) to satisfy (12). The identity preservation parameters which solve (12) are

$$\begin{aligned} \bar{X}_1 &= \bar{X}_0 + V_0 \Delta \bar{t} \\ V_1 &= V_0 \end{aligned} \tag{13}$$

with $\bar{t}_1$, σ_{t1} and C_1 as constraint parameters with known values. Thus the parameters that preserve identity continue motion along a trajectory with the same initial velocity on the same line.

V. EQUIVALENT OBSERVERS

A. Experiment Verification

Acceptance of the outcome of an experiment demands an independent method of verification. The demand can be met in traditional physics where there are implicit assumptions of independent, identifiable, characteristics of the physical system outside the scope of a theory. One example is the Kepler-Newton theory of orbital motion of the planet Mars about the sun, where deviations from the orbit can be observed since the planet is identifiable by characteristics, such as surface features and diameter, outside the scope of the orbital theory. When the theory contains all of the identity information of the physical system, verification cannot be based on system characteristics independent of the theory. An example of such a theory is the bosonic theory of photons, which are indisitnguishable particles fully identified by symmetric wave functions, where no single photon experiment can be repeated using exactly the same photon [8].

Exclusion of verification based on independent system characteristics elevates the significance of verification restricted to statistical analysis of repeatability which underlies the probability densities of the theory. Repeatability can be verified by observers outside the scope of the theory. Thus by keeping the theory incomplete in the sense that it does not bring all observers within the scope of the theories physical descriptions, the requirement of verification with separate system identification moves to a requirement of independent observer verification, especially when the entire physical system description is contained within the theory. Verification of a bosonic photon experimental result never involves preparation of the same photon by a different observer, but instead repetition of the same experimental conditions by an independent observer. Any theory that fully contains the identity of a system cannot support validation based on independent characteristics of the system.

In the theory of special relativity independent observers are represented by Lorentz transformations of space time co-ordinates. Lorentz transformations in one approach arise from quantum consideration [5]. In quantum theory the wave function becomes the object of the transformation through invariant wave equations, associated with corresponding transformations of observables. An information theoretic approach to define the reference frame of equivalent observers has the opportunity to generalize the transformation to another observer to be a transformation of the probability distribution for any type of random variable and theory, without limitation to only space and time quantities. Define an equivalence transformation to be a generalized transformation of the distribution, and therefore of the distribution parameters, to the frame of an equivalent observer. An identity transformation of the probability distribution of a theory defines the local observer and is also an equivalence transformation. A verification capability requires that at least one other observer exist that is not the local observer. A complete and independently verifiable scientific theory has the property of:

1. *Independent Observer Scrutiny*

Every independently verifiable theory must support at least one independent observer other than the local observer, and must necessarily have at least one equivalence transformation that is not the identity transformation.

B. Equivalence Transformations

An equivalent observer must reach the same conclusions from data available to the observer as the local observer of an experiment reaches. Each equivalent observer then has the same information to accept or reject the conclusion of identity preservation under a shift in constraints. The design of the experiment makes the transformation of all initial parameters available to all observers, and execution

of the experiment results in estimates of final parameter values for all observers. An equivalence transformation $\boldsymbol{E}$ transforms parameters P in use by an observer into parameters $P' = \boldsymbol{E}P$ used by an equivalent observer. If $\boldsymbol{E}$ is the equivalent observer transformation, the final discrimination information determined by the equivalent observer is $I\left(\boldsymbol{E}P_1, \boldsymbol{E}P_0\right)$, but the same observer would determine that discrimination as $I\left(\boldsymbol{E}R_1, \boldsymbol{E}P_0\right)$ if identity preservation parameters were observed as expected by the theory, where R_1 is given by (13). The larger the magnitude of the difference between the two discrimination values the more likely is a rejection of the conclusion that identity is preserved by the tested theory, while a very small magnitude difference supports the conclusion. The difference is proportional to the number of bits input into statistical decision methods and for the local observer is equal to

$$\begin{aligned} K\left(P_1, P_0 : \mathrm{A}_\zeta\right) &= I\left(P_1, P_0\right) - I\left(R_1, P_0\right) \\ &= I\left(P_1, P_0\right) - \min_{p \in M\left(P_1 : \mathrm{A}_\zeta(\mathcal{O})\right)} I\left(p, P_0\right) \end{aligned} \quad (14)$$

where the R_1 constraint parameter values $\zeta_1 = \left(\overline{t}_1 \quad \sigma_{t1} \quad C_1\right)$ are also component values in P_1. Since K is input to statistical methods which determine rejection or acceptance of identity preservation, K is invariant:

$$K\left(P_1, P_0 : \mathrm{A}_\zeta\right) = K\left(\boldsymbol{E}P_1, \boldsymbol{E}P_0 : \mathrm{A}_\zeta\right) \quad (15)$$

In the invariant equation (15) the same adjustable constraint A_ζ is used by all equivalent observers of the experiment. Equivalence transformations are dependent on the primary adjustable constraint.

An equivalent observer transformation operates on all parameters $P \in \mathcal{D}$ in a subset $\mathcal{P} \subset \mathcal{Q}$ of the set of parameters $\mathcal{Q}$. As every equivalent observer is also an equivalent observer to any other equivalent observer, the inverse $\boldsymbol{E}^{-1}P$ exists for every P in the range of $\boldsymbol{E}$ and is an equivalence transformation.

An equivalence transformation defines observers in the context of a theory. Equivalence transformations depend on properties of the Primary Constraint of a Confirmation Set. In the Primary Constraint all parameters of the marginal time distribution are constraint parameters; and the equivalence transformations for that adjustable constraint define the transformations which must be associated with all remaining adjustable constraints in the Confirmation Set.

Equivalence transformations are defined by properties, such as invariance, of each transformation. Properties of collections of transformations, such as the verifiable theory property of independent observer scrutiny, also contribute to the definition of equivalence transformations.

C. Equivalent Observer Collections

Define an equivalent observer collection $\mathscr{C}$ as a set containing a set $\mathscr{E} = \left\{\boldsymbol{E}_1, \ldots, \boldsymbol{E}_n, \ldots,\right\}$ of equivalence transformations using primary adjustable constraint A_ζ which is also a member of $\mathscr{C}$, a parameter domain $\mathcal{D} \subseteq \mathcal{Q}$ and range $\mathcal{R} \subseteq \mathcal{Q}$ common to all transformations within $\mathscr{E}$, $\mathscr{C} = \left\{\mathscr{E}, \mathcal{D}, \mathcal{R}, \mathrm{A}_\zeta\right\}$. The properties required of an equivalent observer collection $\mathscr{C}$ are:

1. **Containment Structure**

   Every equivalence transformation $\boldsymbol{E}$ is an element of a set $\mathscr{E}$ contained in some equivalent observer collection $\mathscr{C}$, $\boldsymbol{E} \in \mathscr{E} \in \mathscr{C}$. A theory may produce more than one equivalent observer collection.

2. **Time Postulate Primary Constraint**

   All equivalent observers of an experiment on a probability model of a physical system, with each observer represented by an element $\boldsymbol{E} \in \mathscr{E} \in \mathscr{C}$ choose constraints from an adjustable constraint A_ζ which must be the primary constraint of a confirmation set of adjustable constraints. The time postulate implies that all parameters of the model marginal probability distribution of observation time within A_ζ are constraint parameters.

3. **Invariant Excess Identity Discrimination**

   Every $\boldsymbol{E} \in \mathscr{E}$ preserves identity discrepancy $K\left(P_1, P_0 : \mathrm{A}_\zeta\right)$. Let $S = \left\{\boldsymbol{B}_1, \ldots, \boldsymbol{B}_n, \ldots,\right\}$ be the set of all transformations which preserve K. Then $\forall P_i \in \mathcal{D}, K\left(P_1, P_0 : \mathrm{A}_\zeta\right) = K\left(\boldsymbol{B}_i P_1, \boldsymbol{B}_i P_0 : \mathrm{A}_\zeta\right)$, and $\mathscr{E} \subseteq S$.

4. **Independent Observer Scrutiny**

   Every $\mathscr{E}$ contains at least one $\boldsymbol{E}$ not the identity.

5. **Transitivity**

If $\mathscr{E}' = \left\{\mathscr{E}', \mathcal{D}, \mathcal{R}, \mathrm{A}_\zeta\right\}$ and $\mathscr{E} = \left\{\mathscr{E}, \mathcal{R}, \mathcal{K}, \mathrm{A}_\zeta\right\}$ are equivalent operator collections, then $\mathscr{E}'' = \left\{\mathscr{E}'', \mathcal{D}, \mathcal{K}, \mathrm{A}_\zeta\right\}$ must also be an equivalent operator collection with every equivalence transformation $\boldsymbol{E}'' \in \mathscr{E}''$ equal to the composition $\boldsymbol{E}'' = \boldsymbol{E}'\boldsymbol{E}$ for some $\boldsymbol{E} \in \mathscr{E}$ and $\boldsymbol{E}' \in \mathscr{E}'$, where $\mathcal{K}$ can be any domain or range in an equivalent operator collection. This property implies there must always exist a collection for which the transformation domain equals the range.

6. **Inverse Existence**

   Unique $\boldsymbol{E}^{-1}$ exists for all $\boldsymbol{E} \in \mathscr{E} \in \mathscr{C} = \left\{\mathscr{E}, \mathcal{D}, \mathcal{R}, \mathrm{A}_\zeta\right\}$, and each $\boldsymbol{E}^{-1}$ is a valid equivalence transformation and an element of an equivalent observer collection $\mathscr{C}^{-1} = \left\{\mathscr{E}^{-1}, \mathcal{R}, \mathcal{D}, \mathrm{A}_\zeta\right\}$, where $\mathscr{E}^{-1} = \left\{\boldsymbol{E}^{-1}{}_1, \ldots, \boldsymbol{E}^{-1}{}_n, \ldots\right\}$. Every $\boldsymbol{E}$ is a one to one map.

7. **Maximally Inclusive**

   a. **Maximal Transformation Set**

      Every $\mathscr{E} \in \mathscr{C} = \left\{\mathscr{E}, \mathcal{D}, \mathcal{R}, \mathrm{A}_\zeta\right\}$ contains all possible $\boldsymbol{E}$ that have properties 1 through 6 and produce only valid probability models.

   b. **Maximal Parameter Set**

      Every $\mathcal{D} \in \mathscr{C} = \left\{\mathscr{E}, \mathcal{D}, \mathcal{R}, \mathrm{A}_\zeta\right\}$ and $\mathcal{P} \in \mathscr{C}$ contain all possible parameter values P that give valid probability models for all transformations which satisfy the Maximal Transformation Set property.

These properties define an equivalent observer collection, and the equivalent transformations contained within the collection. There can be more than one observer collection for each probability model.

D. K Invariance

Invariant K is defined by (14), and for the norm density space and time observation model is equal to

$$K = \frac{1}{2}\left(\Delta\bar{X} - V_0\Delta\bar{t}\right)^T C_0{}^{-1}\left(\Delta\bar{X} - V_0\Delta\bar{t}\right) + \frac{1}{2}\sigma_{t1}{}^2\Delta V^T C_0{}^{-1}\Delta V \tag{16}$$

The presence of $\sigma_{t1}{}^2 = \sigma_{t0}{}^2 + \Delta\sigma_t{}^2$ in the second term makes (16) a third order invariant in parameter shifts.

The Equivalent observer collection property of **Invariant Excess Identity Discrimination** requires invariance of K and implies that if $K = 0$ then $K' = K\left(\boldsymbol{E}P_1, \boldsymbol{E}P_0 : \mathrm{A}_\zeta\right) = 0$, so that identity preservation is invariant. With parameter $P = \left(\bar{X}, V, \bar{t}, \sigma_t, C\right)$ set $\boldsymbol{E}P = P' = \left(\bar{X}', V', \bar{t}', \sigma_t', C'\right)$ Identity preservation equation (13) transforms to

$$\begin{aligned} \bar{X}_1' &= \bar{X}_0' + V_0'\Delta\bar{t}' \\ V_1' &= V_0' \end{aligned} \tag{17}$$

Each component of P' potentially depends on components of P. Represent components of parameter P by the invertible transform

$$\begin{aligned} \mu &= \bar{X} - V\bar{t} \quad & \tau &= \bar{t} \quad & \upsilon &= V \\ \eta &= \sigma_t & \varepsilon &= C^{-1} \end{aligned} \tag{18}$$

so that the identity preservation equations (13) and (17) become

$$\begin{aligned} \mu^*{}_1 &= \mu_0 & \mu_1'^* &= \mu_0' \\ \upsilon^*{}_1 &= \upsilon_0 & \upsilon_1'^* &= \upsilon_0' \end{aligned} \tag{19}$$

with $\tau_1, \eta_1, \varepsilon_1$ as constraint parameters and * to indicate identity preservation parameters. In the following analysis assume the transformations are continuous functions of parameters and continuous for all first through third derivatives with respect to components of P_0. Differentiate $\mu_1'^*$ and $\upsilon_1'^*$ in (17) with respect to τ_1, η_1 and ε_1, and use (19) to get

$$\begin{aligned} \frac{\partial}{\partial\tau_1}\upsilon'\left(\mu_0, \upsilon_0, \tau_1, \eta_1, \varepsilon_1\right) &= \frac{\partial}{\partial\tau_1}\upsilon'\left(\mu_0, \upsilon_0, \tau_0, \eta_0, \varepsilon_0\right) = 0 \\ \frac{\partial}{\partial\eta_1}\upsilon'\left(\mu_0, \upsilon_0, \tau_1, \eta_1, \varepsilon_1\right) &= \frac{\partial}{\partial\eta_1}\upsilon'\left(\mu_0, \upsilon_0, \tau_0, \eta_0, \varepsilon_0\right) = 0 \\ \frac{\partial}{\partial\varepsilon_1}\upsilon'\left(\mu_0, \upsilon_0, \tau_1, \eta_1, \varepsilon_1\right) &= \frac{\partial}{\partial\varepsilon_1}\upsilon'\left(\mu_0, \upsilon_0, \tau_0, \eta_0, \varepsilon_0\right) = 0 \\ \frac{\partial}{\partial\tau_1}\mu'\left(\mu_0, \upsilon_0, \tau_1, \eta_1, \varepsilon_1\right) &= \frac{\partial}{\partial\tau_1}\mu'\left(\mu_0, \upsilon_0, \tau_0, \eta_0, \varepsilon_0\right) = 0 \\ \frac{\partial}{\partial\eta_1}\mu'\left(\mu_0, \upsilon_0, \tau_1, \eta_1, \varepsilon_1\right) &= \frac{\partial}{\partial\eta_1}\mu'\left(\mu_0, \upsilon_0, \tau_0, \eta_0, \varepsilon_0\right) = 0 \\ \frac{\partial}{\partial\varepsilon_1}\mu'\left(\mu_0, \upsilon_0, \tau_1, \eta_1, \varepsilon_1\right) &= \frac{\partial}{\partial\varepsilon_1}\mu'\left(\mu_0, \upsilon_0, \tau_0, \eta_0, \varepsilon_0\right) = 0 \end{aligned}$$

which imply that μ' and υ' can be written as

$$\upsilon' = \upsilon'\left(\mu,\upsilon\right) \qquad \mu' = \mu'\left(\mu,\upsilon\right) \tag{20}$$

In the $P = \left(\mu,\upsilon,\tau,\eta,\varepsilon\right)$ representation of parameters the invariance equation $K = K'$ becomes

$$\begin{aligned}&\frac{\eta_1}{2}\Delta\upsilon^T\varepsilon_0\Delta\upsilon + \frac{1}{2}\left(\Delta\mu + \tau_1\Delta\upsilon\right)^T\varepsilon_0\left(\Delta\mu + \tau_1\Delta\upsilon\right) = \\ &\frac{\eta_1'}{2}\Delta\upsilon'^T\varepsilon_0'\Delta\upsilon' + \frac{1}{2}\left(\Delta\mu' + \tau_1'\Delta\upsilon'\right)^T\varepsilon_0'\left(\Delta\mu' + \tau_1'\Delta\upsilon'\right)\end{aligned} \tag{21}$$

Differentiate (21) with respect to η_0 to get

$$\begin{aligned}0 = &\frac{\eta_1'}{2}\Delta\upsilon'^T\frac{\partial\varepsilon_0'}{\partial\eta_0}\Delta\upsilon' \\ &+\frac{1}{2}\left(\Delta\mu' + \tau_1'\Delta\upsilon'\right)^T\frac{\partial\varepsilon_0'}{\partial\eta_0}\left(\Delta\mu' + \tau_1'\Delta\upsilon'\right)\end{aligned} \tag{22}$$

Extraction of valid equivalence transformations $\boldsymbol{E}$ begins with determination of transformations $\boldsymbol{B}$ which keep K invariant, with initial assumption that the domain $\mathcal{D}$ and range $\mathcal{R}$ are the set of all parameters. Then $\mathcal{D}$ and $\mathcal{R}$ are adjusted to satisfy the properties of equivalent observer collections.

Components of P_1' appear in (22) but not components of P_1, so the initial assumption that range and domain equal $\mathcal{D}$ implies P_1' can be any valid parameter in the $\mathcal{D}$. Then $\Delta\mu'$, $\Delta\upsilon'$ and τ_1' can take on an infinity of values.

To establish a procedure that reveals the impact of the infinity of P_1' values that satisfy (22), define unit vectors $\hat{j}_1^T = \begin{pmatrix}1 & 0 & 0\end{pmatrix}$, $\hat{j}_2^T = \begin{pmatrix}0 & 1 & 0\end{pmatrix}$, $\hat{j}_3^T = \begin{pmatrix}0 & 0 & 1\end{pmatrix}$, $\hat{k}_{ij} = \frac{1}{\sqrt{2}}\left(\hat{j}_i + \hat{j}_k\right)$ and $\hat{l}_{ij} = \frac{1}{\sqrt{2}}\left(\hat{j}_i - \hat{j}_j\right)$. If S is any 3x3 symmetric matrix

$$\frac{1}{2}\left(\hat{k}_{ij}^T S\hat{k}_{ij} - \hat{l}_{ij}^T S\hat{l}_{ij}\right) = S_{ik} \tag{23}$$

If $i = k$, (23) becomes $\hat{j}_i^T S\hat{j}_i = S_{ii}$. In (22), set $\Delta\upsilon'$ first to $a\hat{k}_{ij}$ for some scalar a and then to $a\hat{l}_{ij}$ and subtract the equations to replace a $\Delta\upsilon'^T S\Delta\upsilon'$ term with $a^2 S_{ik}$.

Select μ_1 such that $\Delta\mu' = 0$ in (22), and select any small scalar a. Adjust υ_1' to apply the procedure to use (23) for each index pair i, j with indices $1\ldots3$. The result is $0 = \frac{a^2}{2}\left(\eta_1' + \tau_1'^2\right)\frac{\partial\varepsilon_0'}{\partial\eta_0}$, so that $\frac{\partial\varepsilon'}{\partial\eta} = 0$.

Next differentiate (21) with respect to τ_0:

$$\begin{aligned}0 = &\frac{\eta_1'}{2}\Delta\upsilon'^T\frac{\partial\varepsilon_0'}{\partial\tau_0}\Delta\upsilon' \\ &+\frac{1}{2}\left(\Delta\mu' + \tau_1'\Delta\upsilon'\right)^T\frac{\partial\varepsilon_0'}{\partial\tau_0}\left(\Delta\mu' + \tau_1'\Delta\upsilon'\right)\end{aligned} \tag{24}$$

Analysis similar to that for (22) results in $\frac{\partial\varepsilon'}{\partial\tau} = 0$.

Take the second derivative of (21) with respect to η_1:

$$\begin{aligned}0 = &\frac{1}{2}\frac{\partial^2}{\partial{\eta_1}^2}\left(\eta_1' + \tau_1'^2\right)\Delta\upsilon'^T\varepsilon_0'\Delta\upsilon' \\ &+\frac{\partial^2\tau_1'}{\partial{\eta_1}^2}\Delta\upsilon'^T\varepsilon_0'\Delta\mu'\end{aligned} \tag{25}$$

. Components of P_0' appear in (25) but not components of P_0, so P_0' can be any valid parameter. Select $\mu_0' = \mu_1'$ to remove the second term. Since ε_0' is an inverse covariance matrix, $\Delta\upsilon'^T\varepsilon_0'\Delta\upsilon'$ is almost positive definite, and positive for some $\Delta\upsilon'$ so that $\frac{\partial^2}{\partial\eta^2}\left(\eta' + \tau'^2\right) = 0$. Values of $\Delta\mu'$ exist for which $\Delta\upsilon'^T\varepsilon_0'\Delta\mu'$ is not zero, with the result $\frac{\partial^2\eta'}{\partial\eta^2} = \frac{\partial^2\tau'}{\partial\eta^2} = 0$.

The third derivative of (21) with respect to τ_1 is:

$$\begin{aligned}0 = &\frac{1}{2}\frac{\partial^3}{\partial{\tau_1}^3}\left(\eta_1' + \tau_1'^2\right)\Delta\upsilon'^T\varepsilon_0'\Delta\upsilon' \\ &+\frac{\partial^3\tau_1'}{\partial{\tau_1}^3}\Delta\upsilon'^T\varepsilon_0'\Delta\mu'\end{aligned}$$

so that by following a similar analysis to that used for derivatives with respect to η, $\frac{\partial^3}{\partial\tau^3}\left(\eta' + \tau'^2\right) = \frac{\partial^3\tau'}{\partial\tau^3} = \frac{\partial^3\eta'}{\partial\tau^3} = 0$. Presence of a non-zero quadratic terms in the expression integrals violates the unique **Inverse Existence** property, so $\frac{\partial^2\tau'}{\partial\tau^2} = \frac{\partial^2\eta'}{\partial\tau^2} = 0$.

The second derivative of (21) with respect to τ_1 is then:

$$\Delta v^T \varepsilon_0 \Delta v = \left(\frac{\partial \tau_1'}{\partial \tau_1}\right)^2 \Delta v'^T \varepsilon_0' \Delta v' \tag{26}$$

Integrate the η' equations to get $\eta' = a\tau + g\eta + b\tau\eta + h$ where a, b, g, h do not depend on τ or η. Since τ is a mean time with potential unbounded values that can be negative or positive, non-zero a and b allow a negative variance η' and invalid probability model. Thus $a = b = 0$ and $\eta' = g\eta + h$. Integrate the τ' derivatives to obtain $\tau' = q\tau + p\tau\eta + s\eta + r$ where q, p, s, r do not depend on τ or η. Substitute this expression for τ' into (21) and isolate the ${\eta_1}^2$ and ${\eta_1}^2{\tau_1}^2$ terms which appear only on the right transformed side of the equation. The resultant equation requires that $p = s = 0$ and $\tau' = q\tau + r$, with q, r independent of τ or η. A first derivative of (21) by τ_1 results in $\Delta v^T \varepsilon_0 \left(\Delta\mu + \tau_1 \Delta v\right) = \frac{\partial \tau_1'}{\partial \tau_1} \Delta v'^T \varepsilon_0' \left(\Delta\mu' + \tau_1' \Delta v'\right)$ and implies $\frac{\partial \tau'}{\partial \tau} = q \neq 0$.

Using results obtained so far, apply $\frac{\partial^2}{\partial {\mu_1}^2}$ to (21) and then let P_1 approach P_0, $P_1 \to P_0 = P$ to get

$$\begin{aligned} \varepsilon = {\eta'} \frac{\partial {v'}^T}{\partial \mu} \varepsilon' \frac{\partial v'}{\partial \mu} \\ + \left(\frac{\partial \mu'}{\partial \mu} + \tau' \frac{\partial v'}{\partial \mu}\right)^T \varepsilon' \left(\frac{\partial \mu'}{\partial \mu} + \tau' \frac{\partial v'}{\partial \mu}\right) \end{aligned} \tag{27}$$

Differentiate by τ twice to obtain

$$0 = \left(\frac{\partial \tau'}{\partial \tau}\right)^2 \frac{\partial {v'}^T}{\partial \mu} \varepsilon' \frac{\partial v'}{\partial \mu}$$

which shows that $\frac{\partial v'}{\partial \mu} = 0$ and, with (20) that v' is a function of only v, $v' = v'\left(v\right)$.

The derivative of (21) with respect to ε_{ij1} is

$$\begin{aligned} 0 = \frac{1}{2} \frac{\partial \eta_1'}{\partial \varepsilon_{ij1}} \Delta v'^T \varepsilon_0' \Delta v' \\ + \frac{\partial \tau_1'}{\partial \varepsilon_{ij1}} \Delta v'^T \varepsilon_0' \left(\Delta\mu' + \tau_1' \Delta v'\right) \end{aligned} \tag{28}$$

Adjust μ_0' so that $\Delta\mu' = -\tau_1' \Delta v'$ and use the procedure that follows (23) to get $\frac{\partial \eta'}{\partial \varepsilon_{ij}} = 0$. Then select $\Delta\mu' = 0$ to get $\frac{\partial \tau'}{\partial \varepsilon_{ij}} = 0$. Next differentiate (21) by η_1 to get

$$\Delta v^T \varepsilon_0 \Delta v = \frac{\partial \eta_1'}{\partial \eta_1} \Delta v'^T \varepsilon_0' \Delta v' \tag{29}$$

The derivative of (29) by μ_0 is $0 = \frac{\partial \eta_1'}{\partial \eta_1} \Delta v'^T \frac{\partial \varepsilon_0'}{\partial \mu_0} \Delta v'$ and implies $\frac{\partial \varepsilon'}{\partial \mu} = 0$. Differentiation by μ_1 is $0 = \frac{\partial^2 \eta_1'}{\partial \eta_1 \partial \mu_1} \Delta v'^T \varepsilon_0' \Delta v'$. The $\frac{\partial^2 \eta'}{\partial \eta \partial \mu} = 0$ result, combined with the prior $\frac{\partial^2 \eta'}{\partial \eta^2} = 0$ result integrates to $\eta' = g\left(v\right)\eta + h\left(v\right)$ for some positive functions g, h.

Differentiate (21) first by ε_{0ab} and then by ε_{0kl} to get

$$\begin{aligned} 0 = \eta_1' \Delta v'^T \frac{\partial \varepsilon_0'}{\partial \varepsilon_{0ab} \partial \varepsilon_{0kl}} \Delta v' + \\ \left(\Delta\mu' + \tau_1' \Delta v'\right)^T \frac{\partial \varepsilon_0'}{\partial \varepsilon_{0ab} \partial \varepsilon_{0kl}} \left(\Delta\mu' + \tau_1' \Delta v'\right) \end{aligned} \tag{30}$$

With $Q_0 = \frac{\partial \varepsilon_0'}{\partial \varepsilon_{0ab} \partial \varepsilon_{0kl}}$, differentiation of (30) by η_1' as an independent variable yields $\Delta v'^T Q_0 \Delta v' = 0$. Next select $\tau_1' = 0$ so that $\Delta\mu'^T Q_0 \Delta\mu' = 0$. Then (30) becomes $\Delta\mu'^T Q_0 \Delta v' = 0$. Since ${\mu_1}'$ and ${v_1}'$, and thus $\Delta\mu'$ and $\Delta v'$ can be arbitrarily and independently selected, only $Q_0 = 0$ can satisfy the equation for all possible P_1' component values. Thus $\frac{\partial \varepsilon_{ij}'}{\partial \varepsilon_{ab} \partial \varepsilon_{kl}} = 0$, each ε_{ij}' depends linearly on the ε_{ab} components, and the inverse co-variance matrix has the form

$$\varepsilon_{ij}' = T_{ijkl}\left(v\right)\varepsilon_{kl} + D_{ij}\left(v\right) \tag{31}$$

for some T, D. Substitute (31) into (21) and let components of matrix ε_0 approach zero, $\varepsilon_0 \to 0$, or equivalently isolate terms with no ε_{kl} factors, to get

$$\begin{aligned} 0 = &\frac{\eta_1'}{2}\Delta\upsilon'^T D\left(\upsilon_0\right)\Delta\upsilon' \\ &+\frac{1}{2}\left(\Delta\mu' + \tau_1'\Delta\upsilon'\right)^T D\left(\upsilon_0\right)\left(\Delta\mu' + \tau_1'\Delta\upsilon'\right) \end{aligned} \tag{32}$$

Variation of η_1' gives $\Delta\upsilon'^T D\left(\upsilon_0\right)\Delta\upsilon' = 0$, selection of $\tau_1' = 0$ gives $\Delta\mu'^T D\left(\upsilon_0\right)\Delta\mu' = 0$. As μ_1' does not depend on τ_1', then the only term that remains is $\Delta\mu'^T D\left(\upsilon_0\right)\Delta\upsilon' = 0$. $\Delta\mu'$ and $\Delta\upsilon'$ can be selected arbitrarily and independently, so $D\left(\upsilon_0\right) = 0$ and $\varepsilon_{ij}' = T_{ijkl}\left(\upsilon\right)\varepsilon_{kl}$. Any linear transformation of a covariance matrix can take the form $U^T\varepsilon U$ for some orthogonal matrix, so there exists matrix $U\left(\upsilon\right)$ such that

$$\varepsilon' = U\left(\upsilon\right)^T \varepsilon U\left(\upsilon\right) \tag{33}$$

Substitute the derived form for η', $\eta' = g\left(\upsilon\right)\eta + h\left(\upsilon\right)$, into (21) , remove the η_1 factor terms, select $\mu_0{}'$ so that $\Delta\mu = -\tau_1\Delta\upsilon$ to get

$$\begin{aligned} &\frac{h\left(\upsilon\right)}{2}\Delta\upsilon'^T\varepsilon_0'\Delta\upsilon' \\ &+\frac{1}{2}\left(\Delta\mu' + \tau_1'\Delta\upsilon'\right)^T \varepsilon_0'\left(\Delta\mu' + \tau_1'\Delta\upsilon'\right) = 0 \end{aligned} \tag{34}$$

The almost positive definite property of ε_0' in (34) implies $h\left(\upsilon\right) \le 0$. Negative $h\left(\upsilon\right)$ allows negative η' in an improper probability model so that $h\left(\upsilon\right) = 0$ and $\eta' = g\left(\upsilon\right)\eta = \phi\left(\upsilon\right)^2\eta$.

Multiply (29) by η_1 to get

$$\begin{aligned} \frac{1}{2}\eta_1\Delta\upsilon^T\varepsilon_0\Delta\upsilon &= \frac{1}{2}\frac{\partial\eta_1'}{\partial\eta_1}\Delta\upsilon'^T\varepsilon_0'\Delta\upsilon' \\ &= \frac{1}{2}\eta_1\phi\left(\upsilon\right)^2\Delta\upsilon'^T\varepsilon_0'\Delta\upsilon' \\ &= \frac{1}{2}\eta_1{}'\Delta\upsilon'^T\varepsilon_0'\Delta\upsilon' \end{aligned} \tag{35}$$

Take the second derivatives of (35) by components of υ_1 and let $P_1 \to P_0 = P$ to get

$$\varepsilon = \phi\left(\upsilon\right)^2\frac{\partial\upsilon'^T}{\partial\upsilon}\,\varepsilon'\frac{\partial\upsilon'}{\partial\upsilon} \tag{36}$$

Subtract (35) from (21) to get

$$\begin{aligned} &\frac{1}{2}\left(\Delta\mu + \tau_1\Delta\upsilon\right)^T\varepsilon_0\left(\Delta\mu + \tau_1\Delta\upsilon\right) = \\ &\frac{1}{2}\left(\Delta\mu' + \tau_1'\Delta\upsilon'\right)^T\varepsilon_0'\left(\Delta\mu' + \tau_1'\Delta\upsilon'\right) \end{aligned} \tag{37}$$

Let P_1 approach P_0, $P_1 \to P_0 = P$ in the second derivative of (37) with respect to components of μ to get

$$\varepsilon = \frac{\partial\mu'^T}{\partial\mu}\,\varepsilon'\frac{\partial\mu'}{\partial\mu} = \frac{\partial\mu'^T}{\partial\mu}\,U\left(\upsilon\right)^T\varepsilon U\left(\upsilon\right)\frac{\partial\mu'}{\partial\mu} \tag{38}$$

Then

$$\frac{\partial}{\partial\mu}\varepsilon = 2\frac{\partial^2\mu'^T}{\partial\mu^2}\,U\left(\upsilon\right)^T\varepsilon U\left(\upsilon\right)\frac{\partial\mu'}{\partial\mu} = 0$$

so that $\frac{\partial^2\mu'}{\partial\mu^2} = 0$, and μ' is linear in μ, $\mu'\left(\mu,\upsilon\right) = S\left(\upsilon\right)\mu + l\left(\upsilon\right)$ for some matrix $S\left(\upsilon\right)$ and vector $l\left(\upsilon\right)$.

Apply $\frac{\partial^2}{\partial\mu_1{}^2}$ to (37) to get

$$\begin{aligned} \varepsilon_0 = &\frac{\partial^2\tau_1'}{\partial\mu_1{}^2}\Delta\upsilon'^T\varepsilon_0'\left(\Delta\mu' + \tau_1'\Delta\upsilon'\right) \\ &+\left(\frac{\partial\mu_1'}{\partial\mu_1} + \frac{\partial\tau_1'}{\partial\mu_1}\Delta\upsilon'\right)^T\varepsilon_0'\left(\frac{\partial\mu_1'}{\partial\mu_1} + \frac{\partial\tau_1'}{\partial\mu_1}\Delta\upsilon'\right) \end{aligned} \tag{39}$$

Variation $\delta\mu_1'$ gives $\frac{\partial^2\tau'}{\partial\mu^2} = 0$ so that τ' depends linearly on μ.

Since every co-variance inverse co-variance matrix is almost positive definite and symmetric, there exists matrix Z such that $\varepsilon = \mathrm{Z}^T\mathrm{Z}$, and (38) can be written

$$1 = \left(\mathrm{Z}\,U\left(\upsilon\right)\frac{\partial\mu'}{\partial\mu}\mathrm{Z}^{-1}\right)^T\left(\mathrm{Z}\,U\left(\upsilon\right)\frac{\partial\mu'}{\partial\mu}\mathrm{Z}^{-1}\right) \tag{40}$$

Then $\mathrm{Z}\left(\varepsilon\right)U\left(\upsilon\right)\frac{\partial\mu'}{\partial\mu}\mathrm{Z}\left(\varepsilon\right)^{-1} = O$ for some orthogonal matrix O.

Analysis performed to this point implies the equivalence transformation components show the following dependencies:

$$\begin{aligned}
\eta' &= \eta'\left(\eta,\upsilon\right) = \phi\left(\upsilon\right)^2 \eta \\
\mu' &= \mu'\left(\mu,\upsilon\right) = S\left(\upsilon\right)\mu + d\left(\upsilon\right) \\
\varepsilon' &= \varepsilon'\left(\varepsilon,\upsilon\right) = U\left(\upsilon\right)^T \varepsilon U\left(\upsilon\right) \\
\tau' &= \tau'\left(\tau,\upsilon,\mu\right) = \phi\left(\upsilon\right)\tau + l\left(\upsilon\right)\mu + b\left(\upsilon\right) \\
\upsilon' &= \upsilon'\left(\upsilon\right)
\end{aligned} \tag{41}$$

with the sign of ϕ chosen to make $+\phi$ be the τ co-efficient of τ'.

Using (41), equations (38) , (36) become, for some orthogonal matrices R_a, R_b.

$$\begin{aligned}
&\mathrm{Z}\left(\varepsilon\right)U\left(\upsilon\right)S\left(\upsilon\right)\mathrm{Z}\left(\varepsilon\right)^{-1} = R_a \qquad {R_a}^T R_a = 1 \\
&\phi\left(\upsilon\right)\mathrm{Z}\left(\varepsilon\right)U\left(\upsilon\right)\frac{\partial\upsilon'}{\partial\upsilon}\mathrm{Z}\left(\varepsilon\right)^{-1} = R_b \qquad {R_b}^T R_b = 1
\end{aligned} \tag{42}$$

Using index notation with double index summation convention, the first of these equations can be written as

$$\mathrm{Z}_{il}\left(\varepsilon\right)U_{lm}\left(\upsilon\right)S_{mj}\left(\upsilon\right) = R_{ail}\mathrm{Z}_{lj}\left(\varepsilon\right) \tag{43}$$

Any non-singular matrix Z results in a non-singular almost positive definite matrix ε. As there are no limitations on possible values of ε in the model beyond those required of an inverse covariance matrix, Z can be any non-singular 3x3 matrix. Variation of Z_{ij} in (43) shows that

$$U\left(\upsilon\right)S\left(\upsilon\right) = R_a = r1$$

is the product of scalar and the identity matrix. Since ${R_a}^T R_a = 1$, $r_a = \pm 1$. Similar analysis of the second equation of (42) shows that

$$\phi\left(\upsilon\right)U\left(\upsilon\right)\frac{\partial\upsilon'}{\partial\upsilon} = \pm 1$$

and thus

$$\phi\left(\upsilon\right)\frac{\partial\upsilon'}{\partial\upsilon} = \pm S\left(\upsilon\right) \tag{44}.$$

Differentiate (37) by μ_1 and then by μ_0 to get

$$\varepsilon_0 = \left(\frac{\partial\mu_1'}{\partial\mu_1} + \frac{\partial\tau_1'}{\partial\mu_1}\Delta\upsilon'\right)^T \varepsilon_0' \frac{\partial\mu_0'}{\partial\mu_0} \tag{45}$$

so that

$$\frac{\partial\mu_1'}{\partial\mu_1} + \frac{\partial\tau_1'}{\partial\mu_1}\Delta\upsilon'^T = S\left(\upsilon_1\right) + \upsilon_1' l\left(\upsilon_1\right)^T - \upsilon_0' l\left(\upsilon_1\right)^T$$

must not depend on any component of P_1. This requires that l be constant and that for some constant matrix L,

$$\begin{aligned}
S\left(\upsilon\right) &= L - \upsilon' l^T \\
\mu' &= \left(L - \upsilon' l^T\right)\mu + d\left(\upsilon\right) \\
\tau' &= \phi\left(\upsilon\right)\tau + l^T\mu + b\left(\upsilon\right)
\end{aligned}$$

Differentiate (37) by μ_1 , then by τ_1 and finally by υ_1 and re-arrange terms to get

$$\varepsilon_0'^{-1}\varepsilon_0\left(L^T - l{\upsilon_0'}^T\right)^{-1} = \left(\phi_1\frac{\partial\upsilon_1'}{\partial\upsilon_1} + \upsilon_1'\frac{\partial\phi_1}{\partial\upsilon_1}^T - \upsilon_0'\frac{\partial\phi_1}{\partial\upsilon_1}^T\right)$$

The right side of this equation must not depend on components of P_1, which requires that there exist constant vector κ and constant matrix Q such that

$$\frac{\partial\phi}{\partial\upsilon} = -w\kappa \qquad \phi\frac{\partial\upsilon'}{\partial\upsilon} - w\upsilon'\kappa^T = Q$$

with solutions

$$\phi = w\left(1 - \kappa^T\upsilon\right) \qquad \upsilon' = \frac{u\upsilon + j}{w\left(1 - \kappa^T\upsilon\right)} = Q\frac{\upsilon - s}{w\left(1 - \kappa^T\upsilon\right)}$$

with constant scalar w and constant vector $j = -Qs$. And since

$$\phi\left(\upsilon\right)\frac{\partial\upsilon'}{\partial\upsilon} = Q + w\upsilon'\kappa^T = \pm S\left(\upsilon\right) = \pm\left(L - \upsilon' l^T\right)$$

$Q = \pm L,\ l = \mp w\kappa$.

In (37) set $\mu_1 = \mu_0 = 0$ and $\tau_1 = 0$ so that the equation becomes

$$0 = \frac{1}{2}\left(\Delta d + b_1\Delta\upsilon'\right)^T \varepsilon_0'\left(\Delta d + b_1\Delta\upsilon'\right)$$

which requires

$$\begin{aligned}
&\Delta d + b_1\Delta\upsilon' = 0 \\
&d\left(\upsilon_1\right) + b\left(\upsilon_1\right)\upsilon'\left(\upsilon_1\right) = d\left(\upsilon_0\right) + b\left(\upsilon_1\right)\upsilon'\left(\upsilon_0\right)
\end{aligned}$$

Since the right side cannot depend on P_0, $b\left(\upsilon\right) = r$ for some constant scalar r and $d\left(\upsilon\right) = u - r\upsilon'\left(\upsilon\right)$ for some constant vector u.

With these results the τ_1 term of (37) is

$$\begin{aligned}
&\tau_1\Delta\upsilon^T\varepsilon_0\Delta\mu = \\
&\quad\phi\left(\upsilon_1\right)\tau_1\Delta\upsilon'^T\varepsilon_0'\left(\Delta\left(S\left(\upsilon\right)\mu\right) + \left(l^T\mu_1\right)\Delta\upsilon'\right)
\end{aligned}$$

The μ_1 derivative of the τ_1 factors is

$$\begin{aligned}
\Delta\upsilon^T\varepsilon_0 &= \phi\left(\upsilon_1\right)\Delta\upsilon'^T\varepsilon_0'\left(S\left(\upsilon_1\right) + \Delta\upsilon' l^T\right) \\
&= \phi\left(\upsilon_1\right)\Delta\upsilon'^T\varepsilon_0' S\left(\upsilon_0\right)
\end{aligned}$$

and a subsequent υ_1 derivative is

$$\varepsilon_0 = \left(\Delta\upsilon'^T \varepsilon_0' S\left(\upsilon_0\right)\right)\left(\frac{\partial\phi\left(\upsilon_1\right)}{\partial\upsilon_1}\right)^T + \phi\left(\upsilon_1\right)\frac{\partial\upsilon_1'^T}{\partial\upsilon_1}\ \varepsilon_0' S\left(\upsilon_0\right)$$
$$= \left(\Delta\upsilon'^T \varepsilon_0' S\left(\upsilon_0\right)\right)\left(\frac{\partial\phi\left(\upsilon_1\right)}{\partial\upsilon_1}\right)^T + \pm S\left(\upsilon_1\right)^T \varepsilon_0' S\left(\upsilon_0\right)$$

Let P_1 approach P_0 to get $\varepsilon = \pm S\left(\upsilon\right)^T \varepsilon' S\left(\upsilon\right)$. Since ε and ε' are positive definite, only the positive sign is valid and $\varepsilon = S\left(\upsilon\right)^T \varepsilon' S\left(\upsilon\right)$.

These results allow (41) to be written as

$$\begin{aligned}
\eta' &= w^2\left(1-\kappa^T v\right)^2 \eta \\
\upsilon' &= Q\frac{\left(\upsilon - s\right)}{w\left(1-\kappa^T v\right)} \\
\mu' &= Q\left\{1+\frac{\left(\upsilon - s\right)}{\left(1-\kappa^T v\right)}\kappa^T\right\}\mu + u - r\upsilon' \\
\varepsilon' &= \left(Q + w\upsilon'\kappa^T\right)^{T-1}\varepsilon\left(Q + w\upsilon'\kappa^T\right)^{-1} \\
\tau' &= w\left(1-\kappa^T v\right)\tau - w\kappa^T\mu + r
\end{aligned} \tag{46}$$

Restoration of the original parameter representation makes the solution of the invariance equation be

$$\begin{aligned}
\sigma_t'^2 &= w^2\theta\left(V\right)^2\sigma_t^{\ 2} \\
V' &= \boldsymbol{\mathcal{P}}\frac{V-s}{\theta\left(V\right)} \\
\overline{X}' &= w\boldsymbol{\mathcal{P}}\left(\overline{X} - s\overline{t}\right) + u \\
C' &= w^2\left(\boldsymbol{\mathcal{P}} + V'\kappa^T\right)C\left(\boldsymbol{\mathcal{P}} + V'\kappa^T\right)^T \\
\overline{t}' &= w\left(\overline{t} - \kappa^T\overline{X}\right) + r
\end{aligned} \tag{47}$$

with $\boldsymbol{\mathcal{P}}$ defined by $\boldsymbol{\mathcal{P}} = \frac{Q}{w}$ and $\theta\left(V\right) = 1-\kappa^T V$. The transformations given by selection of values of $w, k, s, \boldsymbol{\mathcal{P}}, r$ and u in (47) are transformations that keep K invariant and satisfy the **Invariant Excess Identity Discrimination** property of equivalence transformations under stated assumptions of continuity of derivatives. If $w = \gamma$, $k = -s\gamma c^{-2}$, $\boldsymbol{\mathcal{P}} = 1$, $r = 0$ and $u = 0$ are selected, with $\gamma = \left(1 - s^2c^{-2}\right)^{-\frac{1}{2}}$, then the $\overline{X}'$ and $\overline{t}'$ transformations of (47) correspond to a simple instance of the Lorentz transformation with moving frame velocity s. The set of transformations defined by (47) contains but is larger than the Poincare group. Invariance of K produces not only space and time parameter transformations, but also transformations of velocity, time variance and position given time co-variance from an invariant that is conceptually simpler than the Minkowski space-time separation. To see this consider that for the case $C = \sigma_{c0}^{\ 2}1$, the invariant is

$$K = \frac{1}{2}\sigma_{c0}^{\ -2}\left(\Delta\overline{X} - V_0\Delta\overline{t}\right)^2 + \frac{1}{2}\sigma_{t1}^{\ 2}\sigma_{c0}^{\ -2}\left(\Delta V\right)^2.$$

Analysis leading up to transformation (47) shows that K is the sum of two invariant terms, namely

$$\begin{aligned}
K_X &= \frac{1}{2}\sigma_{c0}^{\ -2}\left(\Delta\overline{X} - V_0\Delta\overline{t}\right)^2 \\
K_V &= \frac{1}{2}\sigma_{t1}^{\ 2}\sigma_{c0}^{\ -2}\left(\Delta V\right)^2
\end{aligned}$$

A Lorentz transformation extended to include transformation of $\sigma_t, \sigma_c,$ and V keeps these terms invariant, and so maintains an elliptic, rather than a hyperbolic invariant, although in a parameter space of up to 21 dimensions.

E. Transformation Selectors

a. Transitivity, Inverse and Composition

$\mathcal{S} = \left(w, \kappa, \boldsymbol{\mathcal{P}}, s, u, r\right)$ is a transformation selector, which picks a single transformation from the set of all transformations defined by (47), and acts as an index of transformations that preserve identity discrepancy. Remaining equivalence transformation properties determine the allowed selectors contained within equivalent observer collections, along with allowed domain and range parameter sets. There is a one to one correspondence between selectors and invariant preserving transformations, so an equivalent observer collection set of transformations $\mathscr{E}$ can be represented by a set of selectors $\mathscr{S}$. This also implies that $\mathscr{C} = \left\{\mathscr{E}, \mathscr{D}, \mathscr{R}, \mathrm{A}_\zeta\right\}$ can be represented by $\mathscr{C} = \left\{\mathscr{S}, \mathscr{D}, \mathscr{R}, \mathrm{A}_\zeta\right\}$.

Exclusion of invalid probability models is implicitly required by the **Maximally Inclusive** properties. In (47), valid probability models require non-zero time variance and non-zero covariance matrix determinant, so that

$$\begin{aligned}
&w \neq 0 \\
&\theta\left(V\right) = 1 - \kappa^T V \neq 0 \\
&\kappa^T V \neq 1
\end{aligned} \tag{48}$$

and

$$\begin{aligned}
&\det\left(\boldsymbol{\mathcal{P}}+V'\kappa^T\right)=\det\left(\boldsymbol{\mathcal{P}}\right)\det\left(\boldsymbol{1}+\frac{V-s}{\theta\left(V\right)}\kappa^T\right)\neq 0\\
&\det\left(\boldsymbol{\mathcal{P}}\right)\neq 0\\
&\det\left(\boldsymbol{1}+\frac{V-s}{\theta\left(V\right)}\kappa^T\right)=1+\kappa^T\frac{V-s}{1-\kappa^T V}=\frac{1-\kappa^T s}{1-\kappa^T V}\neq 0\\
&\kappa^T s\neq 1
\end{aligned}\tag{49}$$

In the selector representation, the **Transitivity** property requires that if $\mathscr{C}=\left\{\mathscr{S},\mathscr{D},\mathscr{R},\mathrm{A}_\zeta\right\}$ and $\mathscr{C}'=\left\{\mathscr{S}',\mathscr{R},\mathscr{K},\mathrm{A}_\zeta\right\}$ are equivalent observer collections, then $\mathscr{C}''=\left\{\mathscr{S}'',\mathscr{D},\mathscr{K},\mathrm{A}_\zeta\right\}$ is an equivalent observer collection with selectors formed from all compositions of $\mathscr{S}'$ and $\mathscr{S}$. This implies that given selector $\mathcal{S}'=\left(w',\kappa',\boldsymbol{\mathcal{P}}',s',u',r'\right)$ with associated invariant transformation $\boldsymbol{E}'$, and selector $\mathcal{S}=\left(w,\kappa,\boldsymbol{\mathcal{P}},s,u,r\right)$ with $\boldsymbol{E}$, the transformation $\boldsymbol{E}'\boldsymbol{E}P=\boldsymbol{E}''P$ for $P\in\mathscr{D}$ must be of the form of (47) with selector $\mathcal{S}''=\left(w'',\kappa'',\boldsymbol{\mathcal{P}}'',s'',u'',r''\right)$. Define $\varepsilon=\boldsymbol{\mathcal{P}}+s'\kappa^T$ and $\lambda=1+\kappa'^T\boldsymbol{\mathcal{P}}s$. Work through the algebra to get

$$\begin{aligned}
w''&=\lambda w'w\\
\kappa''&=\lambda^{-1}\left(\kappa+\boldsymbol{\mathcal{P}}^T\kappa'\right)\\
\boldsymbol{\mathcal{P}}''&=\lambda^{-1}\boldsymbol{\mathcal{P}}'\varepsilon\\
s''&=\varepsilon^{-1}\left(s'+\boldsymbol{\mathcal{P}}s\right)\\
u''&=w'\boldsymbol{\mathcal{P}}'\left(u-rs'\right)+u'\\
r''&=r'+w'\left(r-\kappa'^T u\right)
\end{aligned}\tag{50}$$

as the selector formed by the composition of equivalence transformations.

The **Inverse Existence** Property of equivalent observer collections requires the inverse of an equivalence transformation to be an equivalence transformation within an associated inverse collection. Then there must be a transformation selector for the inverse transformation. Let $\tilde{\boldsymbol{\mathcal{S}}}=\left(\tilde{w},\tilde{\kappa},\tilde{\boldsymbol{\mathcal{P}}},\tilde{s},\tilde{u},\tilde{r}\right)$ represent an equivalence transformation $\boldsymbol{E}$ from domain $\mathscr{D}$ to range $\mathscr{R}$, and let $\boldsymbol{\mathcal{S}}'=\left(w',\kappa',\boldsymbol{\mathcal{P}}',s',u',r'\right)$ represent $\boldsymbol{E}^{-1}$. Using the form (47) for both transformations, and $\boldsymbol{E}^{-1}\boldsymbol{E}=\boldsymbol{1}$, work through the algebra to get

$$\begin{aligned}
\tilde{\varepsilon}&=\tilde{\boldsymbol{\mathcal{P}}}\left(1-\tilde{s}\tilde{\kappa}^T\right)\\
\tilde{\lambda}&=1-\tilde{\kappa}^T\tilde{s}\\
\\
w'&=\tilde{\lambda}^{-1}\tilde{w}^{-1}\\
\kappa'&=-\tilde{\boldsymbol{\mathcal{P}}}^{T-1}\tilde{\kappa}\\
\boldsymbol{\mathcal{P}}'&=\tilde{\lambda}\tilde{\varepsilon}^{-1}\\
s'&=-\tilde{\boldsymbol{\mathcal{P}}}\tilde{s}\\
u'&=-\tilde{w}^{-1}\tilde{\varepsilon}^{-1}\left(\tilde{u}+\tilde{r}\tilde{\boldsymbol{\mathcal{P}}}\tilde{s}\right)\\
r'&=-\tilde{\lambda}^{-1}\tilde{w}^{-1}\left(\tilde{r}+\tilde{\kappa}^T\tilde{\boldsymbol{\mathcal{P}}}^{-1}\tilde{u}\right)
\end{aligned}\tag{51}$$

as the inverse transformation selector. The existence of an inverse (51) requires existence of $\tilde{\boldsymbol{\mathcal{P}}}^{-1}$ so $\det\left(\tilde{\boldsymbol{\mathcal{P}}}\right)\neq 0$.

Next form the composition of a transformation $\boldsymbol{E}$ from $\mathscr{D}\to\mathscr{R}$, represented by $\boldsymbol{\mathcal{S}}=\left(w,\kappa,\boldsymbol{\mathcal{P}},s,u,r\right)$, with the inverse of any transformation $\tilde{\boldsymbol{E}}$ from $\mathscr{D}\to\mathscr{S}$, $\tilde{\boldsymbol{E}}^{-1}$ from $\mathscr{R}\to\mathscr{D}$. The composition $\boldsymbol{E}''=\tilde{\boldsymbol{E}}^{-1}\boldsymbol{E}$ is an equivalence transformation from $\mathscr{D}\to\mathscr{D}$. With $\tilde{\boldsymbol{\mathcal{S}}}=\left(\tilde{w},\tilde{\kappa},\tilde{\boldsymbol{\mathcal{P}}},\tilde{s},\tilde{u},\tilde{r}\right)$ to represent $\tilde{\boldsymbol{E}}$ and $\mathcal{S}''=\left(w'',\kappa'',\boldsymbol{\mathcal{P}}'',s'',u'',r''\right)$ to represent $\boldsymbol{E}''$, use (51) to form the inverse selector and (50) to form the composition selector with result

$$\begin{aligned}
&\tilde{\varepsilon}=\tilde{\boldsymbol{\mathcal{P}}}\left(1-\tilde{s}\tilde{\kappa}^T\right)\qquad\hat{\varepsilon}=\boldsymbol{\mathcal{P}}\left(1-\tilde{s}\kappa^T\right)\\
&\tilde{\lambda}=1-\tilde{\kappa}^T\tilde{s}\qquad\hat{\lambda}=1-\tilde{\kappa}^T s\\
\\
&w''=\hat{\lambda}\tilde{\lambda}^{-1}\tilde{w}^{-1}w\\
&\kappa''=\hat{\lambda}^{-1}\left(\kappa-\tilde{\kappa}\right)\\
&\boldsymbol{\mathcal{P}}''=\hat{\lambda}^{-1}\tilde{\lambda}\tilde{\varepsilon}^{-1}\hat{\varepsilon}\\
&s''=\hat{\varepsilon}^{-1}\boldsymbol{\mathcal{P}}\left(s-\tilde{s}\right)\\
&u''=\tilde{w}^{-1}\tilde{\varepsilon}^{-1}\left(u-\tilde{u}+\left(r-\tilde{r}\right)\tilde{\boldsymbol{\mathcal{P}}}\tilde{s}\right)\\
&r''=\tilde{\lambda}^{-1}\tilde{w}^{-1}\left(r-\tilde{r}+\tilde{\kappa}^T\tilde{\boldsymbol{\mathcal{P}}}^{-1}\left(u-\tilde{u}\right)\right)
\end{aligned}\tag{52}$$

b. Range equals domain case

Consider the case $\mathscr{R}=\mathscr{D}$. Then, given the same adjustable constraint so that all inputs into formation of the observer collection are identical, the **Maximal Transformation Set** property implies $\tilde{\boldsymbol{\mathcal{S}}}$ and $\boldsymbol{\mathcal{S}}$ represent transformations in the same set. Since $\tilde{\boldsymbol{\mathcal{S}}}=\boldsymbol{\mathcal{S}}$ is then possible, making (52) the identity transformation $\boldsymbol{1}$, the identity transformation is always an element of every equivalent observer collection for which range and domain are the same.

In (52) valid probability models require

$$\begin{aligned}\det\left(\boldsymbol{\mathcal{P}}''\right) &= \det\left(\hat{\lambda}^{-1}\tilde{\lambda}\tilde{\varepsilon}^{-1}\hat{\varepsilon}\right) = \frac{\tilde{\lambda}^3\det\left(\hat{\varepsilon}\right)}{\hat{\lambda}^3\det\left(\tilde{\varepsilon}\right)} \\ &= \frac{\tilde{\lambda}^3\det\left(\boldsymbol{\mathcal{P}}\left(1-\tilde{s}\kappa^T\right)\right)}{\hat{\lambda}^3\det\left(\tilde{\boldsymbol{\mathcal{P}}}\left(1-\tilde{s}\tilde{\kappa}^T\right)\right)} = \frac{\tilde{\lambda}^3\left(1-\kappa^T\tilde{s}\right)\det\left(\boldsymbol{\mathcal{P}}\right)}{\hat{\lambda}^3\left(1-\tilde{\kappa}^T\tilde{s}\right)\det\left(\tilde{\boldsymbol{\mathcal{P}}}\right)} \\ &= \frac{\tilde{\lambda}^2\det\left(\boldsymbol{\mathcal{P}}\right)}{\hat{\lambda}^2\det\left(\tilde{\boldsymbol{\mathcal{P}}}\right)} \quad \neq 0,\pm\infty\end{aligned} \tag{53}$$

Since the inverse of (52) must exist, (53) can be neither zero nor infinite. Thus $\tilde{\lambda}\neq 0$ and $\hat{\lambda}\neq 0$, or

$$\begin{aligned}\tilde{\kappa}^T\tilde{s} &\neq 1 \\ \kappa^T\tilde{s} &\neq 1\end{aligned} \tag{54}$$

So if $\mathcal{D}=\mathcal{R}$ then κ and $\tilde{s}$ can be independently chosen from the same set of transformations. This implies that the second inequality in (54) is true for any κ and any s in the transformation set, even if from different transformations and an interdependence exists.

The **Maximal Transformation Set** property implies that all possible κ and s be included in the selectors for the observer collection transformation set. Since there is no imposition of a favored direction, there exist κ and s vectors in all spatial directions in the set. Assume that the magnitude $\left|\kappa\right|$ is unbounded in the set. Given any κ with magnitude at least $\left|s\right|^{-1}$, since s can be in any direction, s can always be chosen close enough to perpendicular to κ that

$$\begin{aligned}\kappa^T s &= \left|\kappa\right|\left|s\right|\cos\theta = \left(\left|s\right|^{-1}+\left|\delta\left|\kappa\right|\right|\right)\left|s\right|\cos\theta \\ &= \left(1+\left|s\right|\left|\delta\left|\kappa\right|\right|\right)\cos\left(\frac{\pi}{2}+\delta\theta\right) = 1\end{aligned} \tag{55}$$

which violates (54). Thus the magnitude of $\left|\kappa\right|$ must be bounded in the transformation set. Switch κ and s in the argument to show that $\left|s\right|$ must also be bounded in a transformation set. Lack of any favored direction implies that the boundaries are spheres. Designate c as the upper bound of $\left|s\right|$ and a as the upper bound of $\left|\kappa\right|$ in a transformation set. Define $\beta_s=\frac{s}{c}$ and $\beta_\kappa=\frac{\kappa}{a}$, with magnitudes less than 1. Then (54) is written

$$\kappa^T s = ac\beta_\kappa{}^T\beta_s \neq 1 \tag{56}$$

with upper bound of $ac=1$. $\kappa^T s$ can take any values with magnitude less than one, including values very near one, and the largest value ac must therefore equal 1:

$$\begin{aligned}a &= \frac{1}{c} \\ \kappa^T s &= \beta_\kappa{}^T\beta_s\end{aligned} \tag{57}$$

Reasoning similar to that used in (54) through (57) can be applied to (48), $\kappa^T V\neq 1$. If the magnitude of V is $\left|V\right|>c$, and κ of any magnitude $\left|\kappa\right|<\frac{1}{c}$, the angle formed by κ and V can be chosen close enough to $\frac{\pi}{2}$ radians to make the product equal 1. Thus the magnitude of V is bounded by c:

$$\left|V\right|<c \tag{58}$$

Different values of c result in different parameter domain sets and therefore different equivalent observer collections. Transformations to a range not the same as the domain are to observers which use different values of c. The value of c acts as an index selector for equivalent observer collections, and is a meta selector, providing data about collections of transformations that primarily results from equivalence transformation properties other than invariance.

c. Transformation of velocity bound

Assume $\mathcal{R}=\mathcal{D}$, so both $\left|V\right|$ and $\left|V'\right|$ are bounded by c. Define $\beta=\frac{V}{c}$ and $\beta'=\frac{V'}{c}$ so $\left|\beta\right|<1$ and $\left|\beta'\right|<1$. The transformation of velocity in (47) is, in ς,β,β_r and c terms,

$$\beta' = \boldsymbol{\mathcal{P}}\frac{\beta-\beta_s}{1-\beta_\kappa{}^T\beta} \tag{59}$$

Let r be the magnitude of $\beta-\beta_s$, and $\hat{q}$ be the unit vector in the direction of $\beta-\beta_s$, so $\beta-\beta_s=r\hat{q}$ with $r\geq 0$. Define the square magnitude of the transformed normalized velocity β' as a function $\Pi\left(\beta\right)=\beta'^T\beta'$ so that $\Pi\left(\beta\right)=r^2\theta^{-2}R$ with $R=\hat{q}^T\boldsymbol{\mathcal{P}}^T\boldsymbol{\mathcal{P}}\hat{q}$ and $\theta=1-\beta_\kappa{}^T\beta=1-\beta_\kappa{}^T\beta_s-r\beta_\kappa{}^T\hat{q}$. The minimum value of Π is $\Pi=0$, which occurs at $\beta=\beta_s$, a stationary point of Π. Fix $\hat{q}$ so that the derivative

$$\frac{d}{dr}\Pi(\beta)=\frac{d}{dr}\left(r^2\theta^{-2}\right)R=2\left(r\theta^{-1}\right)\left(\theta^{-1}-r\theta^{-2}\frac{d\theta}{dr}\right)R$$
$$=2\left(r\theta^{-1}\right)\left(\theta^{-1}+r\theta^{-2}\beta_\kappa{}^T\hat{q}\right)R$$
$$=2\left(r\theta^{-3}\right)\left(\theta+r\beta_\kappa{}^T\hat{q}\right)R=2\left(r\theta^{-3}\right)\left(1-\beta_\kappa{}^T\beta_s\right)R$$

Since $R>0$, $r>0$, $\beta_\kappa{}^T\beta_s<1$, and $\beta_\kappa{}^T\beta$, the derivative $\frac{d}{dr}\Pi(\beta)>0$ for all possible directions $\hat{q}$. Thus there can be no other stationary points besides $\beta=\beta_s$.

Then the maximum value of Π, which must equal 1 if $\mathcal{R}=\mathcal{D}$, must occur for β with velocity at the boundary of $\mathcal{D}$, where also $\beta^T\beta=1$ is a maximum. Therefore the boundary of $\mathcal{D}$ maps into the boundary of $\mathcal{R}$, which by assumption $\mathcal{R}=\mathcal{D}$, is also the boundary of $\mathcal{D}$, and (59) must map all magnitude one β to magnitude one β'.

Every range $\mathcal{R}$ can also be a domain, so if $\mathcal{R}\neq\mathcal{D}$, $\mathcal{R}$ and $\mathcal{D}$ are characterized by different values of the velocity bound (58). Designate the bound for $\mathcal{D}$ by c and for $\mathcal{R}$ by c'. Define $\zeta=\frac{c'}{c}$. Then (59) becomes

$$\frac{V'}{c'}=\frac{V'}{\zeta c}=\frac{\beta'}{\zeta}=\boldsymbol{\mathcal{P}}\frac{\beta-\beta_s}{1-\beta_\kappa{}^T\beta} \tag{60}$$

The maximum value of the magnitude right side of (60) is the maximum value of Π, which equals 1, so the maximum value of the magnitude of β' is ζ. Since the maximum magnitude of Π occurs at the boundary of $\mathcal{D}$, the boundary of $\mathcal{D}$ maps into the boundary of $\mathcal{R}$.

F. Lorentz Transformation

Since the analysis of the domain structure shows that an equivalence transformation maps any velocity on the boundary of $\mathcal{D}$, $V=c\Upsilon$, with unit vector Υ, to a velocity also magnitude $c'=\zeta c$ on the boundary of $\mathcal{R}$.

$$V'=\boldsymbol{\mathcal{P}}\frac{V-s}{1-\kappa^{\mathrm{T}}V}=\boldsymbol{\mathcal{P}}\frac{c\Upsilon-s}{1-\kappa^{\mathrm{T}}V}=\zeta c\Upsilon' \tag{61}$$

for some unit vector Υ'. From this equation form the squared magnitude of the transformed velocity as

$$\left(c\Upsilon-s\right)^{\mathrm{T}}\frac{\boldsymbol{\mathcal{P}}^{\mathrm{T}}\boldsymbol{\mathcal{P}}}{\left(1-\beta_\kappa{}^{\mathrm{T}}\Upsilon\right)^2}\left(c\Upsilon-s\right)=\zeta^2c^2$$

which can be written as

$$\left(1-\beta_\kappa{}^{\mathrm{T}}\Upsilon\right)^2=\left(\Upsilon-\beta_s\right)^{\mathrm{T}}\boldsymbol{\mathcal{A}}\left(\Upsilon-\beta_s\right) \tag{62}$$

with $\boldsymbol{\mathcal{A}}=\boldsymbol{\mathcal{S}}^{\mathrm{T}}\boldsymbol{\mathcal{S}}$, and $\boldsymbol{\mathcal{S}}=\zeta^{-1}\boldsymbol{\mathcal{P}}$. Since (61) restricts only the velocity magnitude and not direction, equation (62) must be true for all directions of unit vector Υ. Reverse the direction of the unit vector to obtain $\left(1+\beta_\kappa{}^{\mathrm{T}}\Upsilon\right)^2=\left(\Upsilon+\beta_s\right)^{\mathrm{T}}\boldsymbol{\mathcal{A}}\left(\Upsilon+\beta_s\right)$. These two equations imply that

$$\Upsilon^{\mathrm{T}}\left(\boldsymbol{\mathcal{A}}-\beta_\kappa\beta_\kappa{}^{\mathrm{T}}-1\left(1-\beta_s{}^{\mathrm{T}}\boldsymbol{\mathcal{A}}\beta_s\right)\right)\Upsilon=0$$
$$\left(\beta_\kappa{}^{\mathrm{T}}-\beta_s{}^{\mathrm{T}}\boldsymbol{\mathcal{A}}\right)\Upsilon=0$$

Since these equations must be true for any unit vector Υ,

$$\boldsymbol{\mathcal{A}}=\beta_\kappa\beta_\kappa{}^{\mathrm{T}}+1\left(1-\beta_s{}^{\mathrm{T}}\boldsymbol{\mathcal{A}}\beta_s\right)$$
$$\beta_\kappa=\boldsymbol{\mathcal{A}}\beta_s=\left(\beta_\kappa\beta_\kappa{}^{\mathrm{T}}+1\left(1-\beta_s{}^{\mathrm{T}}\boldsymbol{\mathcal{A}}\beta_s\right)\right)\beta_s \tag{63}$$

Re-arrange the second equation to get

$$\beta_\kappa\left(1-\beta_\kappa{}^{\mathrm{T}}\beta_s\right)=\beta_\kappa\left(1-\beta_s{}^{\mathrm{T}}\boldsymbol{\mathcal{A}}\beta_s\right)$$
$$=\left(1-\beta_s{}^{\mathrm{T}}\boldsymbol{\mathcal{A}}\beta_s\right)\beta_s \tag{64}$$

If the expressions in parenthesis in (63) were to vanish, then $\boldsymbol{\mathcal{A}}=\beta_\kappa\beta_\kappa{}^{\mathrm{T}}$ for which $\det(\boldsymbol{\mathcal{A}})=0$ and therefore, $\det(\boldsymbol{\mathcal{P}})=0$, in violation of the **Inverse Existence** Property. Thus the terms in parenthesis do not vanish. Consequently (64) requires $\beta_\kappa=\beta_s$, or $\kappa=\frac{s}{c^2}$, and β_s is an eigenvector of $\boldsymbol{\mathcal{A}}$. The definition of $\boldsymbol{\mathcal{A}}$ makes $\boldsymbol{\mathcal{A}}$ a positive definite matrix, so it is possible to write $\boldsymbol{\mathcal{A}}=\boldsymbol{\mathcal{M}}^{\mathrm{T}}\boldsymbol{\mathcal{D}}\boldsymbol{\mathcal{M}}$ for some diagonal matrix $\boldsymbol{\mathcal{D}}$ and some orthogonal matrix $\boldsymbol{\mathcal{M}}$, $\boldsymbol{\mathcal{M}}^{\mathrm{T}}\boldsymbol{\mathcal{M}}=\boldsymbol{\mathcal{M}}\boldsymbol{\mathcal{M}}^{\mathrm{T}}=1$. With $\boldsymbol{\mathcal{A}}$ expressed in this manner the first line of (63) becomes

$$\boldsymbol{\mathcal{D}}=\boldsymbol{\mathcal{M}}\left(\beta_s\beta_s{}^{\mathrm{T}}+1\left(1-\beta_s{}^{\mathrm{T}}\boldsymbol{\mathcal{A}}\beta_s\right)\right)\boldsymbol{\mathcal{M}}^{\mathrm{T}}$$
$$=\left(\boldsymbol{\mathcal{M}}\beta_s\right)\left(\boldsymbol{\mathcal{M}}\beta_s\right)^{\mathrm{T}}+1\left(1-\beta_s{}^{\mathrm{T}}\beta_s\right) \tag{65}$$

Define $\gamma=\left(1-\beta_s{}^{\mathrm{T}}\beta_s\right)^{-\frac{1}{2}}$. Then the component equations of (65) are

$$\boldsymbol{\mathcal{D}}_{ii}\delta_{ij}=\left(\boldsymbol{\mathcal{M}}\beta_s\right)_i\left(\boldsymbol{\mathcal{M}}\beta_s\right)_j+\gamma^{-2}\delta_{ij}$$

which imply that only one of the three $\left(\boldsymbol{\mathcal{M}}\beta_s\right)_i$ can be non-zero. Select the index of the non-zero component of $\boldsymbol{\mathcal{M}}\beta_s$ as $i=1$. In the following use of either of the other two axis for the non-zero component yields the same $\boldsymbol{\mathcal{A}}$. Since the matrix $\boldsymbol{\mathcal{M}}$ is orthogonal, $\left|\boldsymbol{\mathcal{M}}\beta_s\right|=\left|\beta_s\right|=\left|\beta_\kappa\right|$, $\left(\boldsymbol{\mathcal{M}}\beta_s\right)_1^2=\beta_s{}^{\mathrm{T}}\beta_s$ and

$$\boldsymbol{\mathcal{D}}_{11}=\beta_s{}^{\mathrm{T}}\beta_s+\gamma^{-2}=1\qquad\boldsymbol{\mathcal{D}}_{22}=\boldsymbol{\mathcal{D}}_{33}=\gamma^{-2}$$

Selection of any other index as the non-zero component results in a corresponding permutation of the $\mathcal{D}$ indices. Matrix $\mathcal{M}$ rotates β_s to align with a coordinate axis. The definition of $\mathcal{A}$ implies

$$\mathcal{A} = \mathcal{S}^{\mathrm{T}}\mathcal{S} = \mathcal{M}^{\mathrm{T}}\mathcal{D}\mathcal{M} = \left(\Lambda\mathcal{M}\right)^{\mathrm{T}}\left(\Lambda\mathcal{M}\right)$$

with $\Lambda_{ij} = \mathcal{D}_{ij}^{\frac{1}{2}}$. Consequently, $\mathcal{S} = \tilde{\mathcal{O}}\Lambda\mathcal{M}$ for some orthogonal matrix $\tilde{\mathcal{O}}$, $\tilde{\mathcal{O}}^{\mathrm{T}}\tilde{\mathcal{O}} = 1$.

Define a position unit scale factor d by reference to an axis perpendicular to β_s as

$$d = w\zeta\left(\tilde{\mathcal{O}}^{\mathrm{T}}\mathcal{S}\mathcal{M}^{\mathrm{T}}\right)_{33} = w\zeta\Lambda_{33} = w\zeta\gamma^{-1}$$

Selector w is then given by

$$w = \left(\frac{d}{\zeta}\right)\gamma$$

With d as position scale factor and ζ as velocity scale factor, the equivalence transformation (47) becomes

$$\begin{aligned}
\bar{X}' &= d\gamma\mathcal{S}\left(\bar{X} - s\bar{t}\right) + u \\
\bar{t}' &= \frac{d}{\zeta}\gamma\left(\bar{t} - \frac{s^T\bar{X}}{c^2}\right) + r \\
V' &= \zeta\mathcal{S}\frac{V - s}{\left(1 - \frac{s^TV}{c^2}\right)} \\
\sigma_t'^2 &= \left(\frac{d}{\zeta}\right)^2\gamma^2\left(1 - \frac{s^T}{c^2}V\right)^2\sigma_t^{\,2} \\
C' &= d^2\gamma^2\mathcal{S}\left(\mathcal{I} + \frac{\left(V - s\right)s^T}{c^2 - s^TV}\right)C\left(\mathcal{I} + \frac{s\left(V - s\right)^T}{c^2 - s^TV}\right)\mathcal{S}^T
\end{aligned} \tag{66}$$

with $\mathcal{S} = \tilde{\mathcal{O}}\begin{pmatrix}1 & 0 & 0\\ 0 & \gamma^{-1} & 0\\ 0 & 0 & \gamma^{-1}\end{pmatrix}\mathcal{M}$, and $\gamma = \left(1 - \frac{s^{\mathrm{T}}s}{c^2}\right)^{-\frac{1}{2}}$, for some orthogonal matrices $\tilde{\mathcal{O}}$ and $\mathcal{M}$, such that matrix $\mathcal{M}$ rotates s into a vector which lies on the x axis $\mathcal{M}s = +\left|s\right|\hat{x}$. Define $\mathcal{O} = \tilde{\mathcal{O}}\mathcal{M}$ so that $\tilde{\mathcal{O}} = \mathcal{O}\mathcal{M}^T$. Then $\mathcal{S} = \mathcal{O}\mathcal{M}^T\begin{pmatrix}1 & 0 & 0\\ 0 & \gamma^{-1} & 0\\ 0 & 0 & \gamma^{-1}\end{pmatrix}\mathcal{M} = \mathcal{O}\mathcal{W}$ with $\mathcal{W}$ defined by $\mathcal{W} = \mathcal{M}^T\begin{pmatrix}1 & 0 & 0\\ 0 & \gamma^{-1} & 0\\ 0 & 0 & \gamma^{-1}\end{pmatrix}\mathcal{M}$. $\mathcal{W}$ first rotates s to lie along the x axis, next multiplies the components of V perpendicular to s by γ^{-1} and then returns s to the original orientation. The matrix $\mathcal{M}$ which rotates V to the x axis is

$$\mathcal{M}v = \left|s\right|\hat{x} \quad \left|s\right| = \sqrt{s_x^{\,2} + s_y^{\,2} + s_z^{\,2}} \quad \left|s_\perp\right| = \sqrt{s_y^{\,2} + s_z^{\,2}}$$

$$\mathcal{M} = \begin{pmatrix}\frac{s_x}{\left|s\right|} & \frac{s_y}{\left|s\right|} & \frac{s_z}{\left|s\right|}\\ \frac{\left|s_\perp\right|}{\left|s\right|} & \frac{-s_xs_y}{\left|s\right|\left|s_\perp\right|} & \frac{-s_xs_z}{\left|s\right|\left|s_\perp\right|}\\ 0 & \frac{s_z}{\left|s\right|\left|s_\perp\right|} & \frac{-s_y}{\left|s\right|\left|s_\perp\right|}\end{pmatrix} \quad \Gamma^{-1} = \begin{pmatrix}1 & 0 & 0\\ 0 & \gamma^{-1} & 0\\ 0 & 0 & \gamma^{-1}\end{pmatrix}$$

$$\mathcal{W} = \mathcal{M}^T\Gamma^{-1}\mathcal{M} = \quad \gamma^{-1}1 + (1 - \gamma^{-1})\frac{ss^T}{s^Ts}$$

Application of $\mathcal{W}$ to the velocity transformation term results in the traditional relativistic velocity transformation. $\mathcal{W}$ is the same no matter which axis the frame velocity is first rotated into, and the matrix can be expressed as

$$\mathcal{W} = \gamma^{-1}1 + (1 - \gamma^{-1})\frac{ss^T}{s^Ts}$$

With these results the equivalence transformation (66) is

$$\begin{aligned}
\bar{X}' &= d\mathcal{O}\left(1 + (\gamma - 1)\frac{ss^T}{s^Ts}\right)\left(\bar{X} - s\bar{t}\right) + q \\
\bar{t}' &= \left(\frac{d}{\zeta}\right)\gamma\left(\bar{t} - \frac{s^T\bar{X}}{c^2}\right) + r \\
V' &= \zeta\gamma^{-1}\mathcal{O}\left(1 + (\gamma - 1)\frac{ss^T}{s^Ts}\right)\frac{V - s}{\left(1 - \frac{s^TV}{c^2}\right)} \\
\sigma_t'^2 &= \left(\frac{d}{\zeta}\right)^2\gamma^2\left(1 - \frac{s^TV}{c^2}\right)^2\sigma_t^{\,2} \\
C' &= d^2\mathcal{O}\left(1 + (\gamma - 1)\frac{ss^T}{s^Ts}\right) \\
&\quad \left(\mathcal{I} + \frac{\left(V - s\right)s^{\mathrm{T}}}{c^2 - s^TV}\right)C\left(\mathcal{I} + \frac{s\left(V - s\right)^{\mathrm{T}}}{c^2 - s^TV}\right) \\
&\quad \left(1 + (\gamma - 1)\frac{ss^T}{s^Ts}\right)\mathcal{O}^{\mathrm{T}}
\end{aligned} \tag{67}$$

The first three equations of (67) are a combination of time and velocity units scaling transformations and Lorentz transformations in time and three spatial dimensions [9]. These are relativistic transformations which account for the possibility of different measurement scales. In addition to conventional relativistic Lorentz transformations, the equivalence transformation defines precisely the transformation of the observation time variance and the spatial covariance given time.

G. Small Parameter Shift

Consider a general continuous probability model with continuous derivatives, vector R of real scalar responsive parameters, and vector W of real scalar constraint parameters. Designate P as the vector of parameters with components from X, V and W. Under sufficiently small $\Delta P_i = P_{1i} - P_{0i}$ the discrimination information [10] is, to second order in ΔP,

$$I\left(P_1, P_0\right) \approx \frac{1}{2} \Delta P^T \boldsymbol{\mathcal{F}} \Delta P \tag{68}$$

where $\boldsymbol{\mathcal{F}}$ is the Fisher information matrix defined by

$$\boldsymbol{\mathcal{F}}_{ij} = \int_{\forall R} dv_R \; f\left(R : P_0\right) \frac{\partial \ln f\left(R : P_0\right)}{\partial P_{0i}} \frac{\partial \ln f\left(R : P_0\right)}{\partial P_{0j}}$$

$\boldsymbol{\mathcal{F}}$ in block matrix form is

$$\boldsymbol{\mathcal{F}} = \begin{pmatrix} \boldsymbol{\mathcal{F}}_{RR} & \boldsymbol{\mathcal{F}}_{RW} \\ \boldsymbol{\mathcal{F}}_{WR} & \boldsymbol{\mathcal{F}}_{WW} \end{pmatrix}$$

which makes (68) become

$$\begin{aligned} I\left(P_1, P_0\right) \approx & \frac{1}{2} \Delta R^T \boldsymbol{\mathcal{F}}_{RR} \Delta R + \Delta R^T \boldsymbol{\mathcal{F}}_{RW} \Delta W \\ & + \frac{1}{2} \Delta W^T \boldsymbol{\mathcal{F}}_{WW} \Delta W \end{aligned} \tag{69}$$

The variational equation

$$\frac{\partial}{\partial R_1} I\left(P_1, P_0\right) = 0$$

defines the stationary point, and minimum, with solution R_1^*:

$$\Delta R^* = R_1^* - R_0 = -\boldsymbol{\mathcal{F}}_{RR}{}^{-1} \boldsymbol{\mathcal{F}}_{RW} \Delta W \tag{70}$$

Substitution into (69) with result

$$\begin{aligned} I\left(P_1^*, P_0\right) \approx & \frac{1}{2} \Delta W^T \boldsymbol{\mathcal{F}}_{RW} \boldsymbol{\mathcal{F}}_{RR}{}^{-1} \boldsymbol{\mathcal{F}}_{RW} \Delta W \\ & - \Delta W^T \boldsymbol{\mathcal{F}}_{RW} \boldsymbol{\mathcal{F}}_{RR}{}^{-1} \boldsymbol{\mathcal{F}}_{RW} \Delta W \\ & + \frac{1}{2} \Delta W^T \boldsymbol{\mathcal{F}}_{WW} \Delta W \end{aligned} \tag{71}$$

Now the excess discrimination invariant is

$$\begin{aligned} K\left(P_1, P_0\right) &= I\left(P_1, P_0\right) - I\left(P_1^*, P_0\right) \\ &= \frac{1}{2} \left(\Delta R + \boldsymbol{\mathcal{F}}_{RR}{}^{-1} \boldsymbol{\mathcal{F}}_{RW} \Delta W\right)^T \boldsymbol{\mathcal{F}}_{RR} \\ &\quad \left(\Delta R + \boldsymbol{\mathcal{F}}_{RR}{}^{-1} \boldsymbol{\mathcal{F}}_{RW} \Delta W\right) \end{aligned} \tag{72}$$

In the simple constant velocity probability model of (10) the response, or un-constrained component vector R is composed of components of $\bar{X}$ and V, while W is composed of components of $\bar{t}, \sigma_t, C$. The second order excess discrimination (72) for the constant velocity model is

$$\begin{aligned} K\left(P_1, P_0\right) = & \frac{1}{2} \left(\Delta \bar{X} - V_0 \Delta \bar{t}\right)^T C_0{}^{-1} \left(\Delta \bar{X} - V_0 \Delta \bar{t}\right) \\ & + \frac{1}{2} \sigma_{t0}{}^2 \Delta V^T C_0{}^{-1} \Delta V \end{aligned} \tag{73}$$

The correct value of K given by (16) differs from (73) by replacement of σ_{t0} with σ_{t1}. Since $\sigma_{t1} / \sigma_{t0}$ is not invariant under the equivalence transformation, the transformation does not exactly keep (73) invariant, but only to a second order approximation. The discrepancy also implies that the transformation that preserves invariance of (73) is only an approximation to the correct transformation. Extension of (68) to third order terms recovers (16). Despite initial appearances (16) is actually a third order invariant since $\sigma_{t1}{}^2 = \sigma_{t0}{}^2 + \Delta \sigma_t{}^2$.

H. Equivalence Transformations Recap

Equivalence transformations are defined on parameters of a probability model by properties which include invariance of information in a deviation from parameter values expected after a controlled parameter component shift. The invariant information can be used to determine if parameter estimates confirm to expected system behavior, so equivalence transformations keep invariant related statistical decision processes.

Remaining properties of equivalence transformations define relations between the transformations and structures of transformation domain and range needed to generate valid probability models for all elements of the domain. Some of the significant concepts, properties and mathematical tools developed to support them are

1. The property of independent observer scrutiny requires that collections of equivalence transformations contain transformations other than identity. Independent observers are needed to confirm a theory that does not support identification of the observed system independent of the applied model.

2. Adjustable parameter constraints allow parameter shifts to be defined in an indexed set theory context. A parameter value acts as index to an adjustable constraint set element, which is a constraint set and the value of a constraint.

3. The time postulate requires that observation time parameters be constraints in an equivalent observer collection. Under the postulate time as control is the essential characteristic which distinguishes time from other quantities.

Statistical samples on a high correlation coefficient probability model show that a multi-variate normal density in observed position and time random variables is a reasonable representation of uniform motion. Application of the properties of equivalence transformations to this probability model results in equivalence transformations which are Lorentz transformations of mean observed position and time, Lorentz transformation of velocity, and specific transformations of the co-variance matrix and standard deviation of observation times. The elliptic information invariant is conceptually simpler than the Minkowski space-time interval, though in higher dimension than four since there are more parameter components. Structural properties on parameter domain and range require that velocity magnitude be bounded, without introduction of any external light speed concept.

Works Cited

[1] J. Hauth, "Grey-Box Modelling for Nonlinear Systems," *Dissertation im Fachbereich Mathematik der Technischen,* December 2008.

[2] K. V. Chernyshov, "An Information Theoretic Approach to System Identification via Input/Output Signal Processing," in *SPECOM'2006*, St. Petersburg, Russia, 2006.

[3] K. Petr and J. Vlcek, "On Improving Probability Models in System Identification," *Mathematical Modeling,* vol. 7, no. 1, pp. 182-186, 2012.

[4] S. Kullback, Information Theory and Statistics, Mineola, NY: Dover Publications, 1997, p. 399.

[5] P. A. Hoehn and M. p. Mueller, "An operational approach to spacetime symmetries:Lorentz transformations from quantum communication," 1 February 2016. [Online]. Available: http://arxiv.org/abs/1412.8462v4.

[6] O. Certik, "Simple Derivation of the special relativity without the speed of light axiom," 7 10 2007. [Online]. Available: http://arxiv.org/pdf/0710.3398.pdf.

[7] A. Pelissetto and T. Massimo, "Getting the Lorentz transformations without requiring an invariant speed," 2015.

[8] B. Braverman, S. Ravets, S. Kocsis, M. J. Stevens, R. P. Mirin, L. K. Shalm1 and A. M. Steinberg, "Observing the Average Trajectories of Single Photons in a Two-Slit Interferometer," *Science,* vol. 332, no. 6034, pp. 1170-1173, 3 June 2011.

[9] "Lorentz transformations in three dimensions," 22 6 2011. [Online]. Available: http://www.physicspages.com/2011/06/22/lorentz-transformations-in-three-dimensions/.

[10] S. Kullback, "Fisher's Information," in *Information Theory and Statistics*, Mineola, NY: Dover, 1997, pp. 26-28.