Bogdan Rusu
Department of Philosophy
University "Dunărea de Jos" Galați
Romania

ON FORMAL CONCEPTS OF COLOR-INVERTED EARTH

Abstract

This paper deals with the problem of color inversion, in the form of the intra-personal inversion. Its main goal is to show that there is something mathematically wrong with any concept of color-inverted Earth. Two such concepts (according to two different ways of conceiving color-inversion) are formally sketched and investigated. An intuitive example is produced in order to show that color-inversion is not distributive to color composition, this mathematical property of inversion being the main assumption of any scenario of intra-personal spectral inversion conceivable.

Key words: color-inversion, color-composition, qualia, spectral inversion

Since Shoemaker's original "Inverted Spectrum" (1982) and Block's "Inverted Earth" (1990) there has been a continuous debate concerning the possibility of intra-subjective and/or extra-subjective spectral inversion. It is enough to notice that two of the most important books published in the last decades in the field of the philosophy of mind, namely Dennett's *Consciousness Explained* (1991) and Chalmers' *The Conscious Mind* (1996), give minute analysis to the scenario of spectral inversion, to grasp the importance of the subject. These scenarios make useful counter-arguments to functionalism and representationalism about qualia. More, they seem to be empirically realizable with virtual reality goggles, so there are good reasons to discuss them.

In this paper my aim is a modest one: to contest the possibility of the Inverted Earth scenario, without getting into the issue of the metaphysical nature of qualia, or of the nature of perception. In order to do that, I will do the following: In section 1 I will give a brief account of the two thought-experiments. Then, in section 2, I will examine the notion of "color inversion" and look at the possible ways of assigning physical and mathematical meaning to the concept. In the next section, I will define the notion of "formal concept of color inverted Earth" and mention two such formal concepts, based on the different meanings of "color inversion" discerned in the previous section. The fourth section will be concerned with the tentative of mathematically treating the important notions involved in these concepts. Then, in the following section, I will appeal to an "intuition pump", which, in connection with the mathematical analysis of the problem, will lead to a possible way of proving the impossibility of the Inverted Earth scenario. The final section will consist in a few remarks about the strength of the argument.

1. The scenario of inverted spectrum[1] presents the aspect of an interpersonal inversion of qualia. Apparently, the problem was first raised by Locke, but its most quoted contemporary formulation is Shoemaker's. This mental experiment is based on the intuition that is very plausible that things might look to someone exactly inversely than they look to me. For example, he might see in stead of each color of the spectrum the complementary color: where I

[1] Shoemaker, "The Inverted Spectrum", *Journal of Philosophy,* 79: 357-81

see green, he could see red and vice versa; where I see blue, he could see orange and vice versa, etc. Due to the private character of our color experiences, that means of their being essentially connected with a first-person point of view, I could never realize that the other person perceives colors exactly inversely than I do. This person could operate the same discriminations which I operate, and, due to the public character of the language, use the words referring to colors exactly as I use them.

The complication of this experiment by making the spectral inversion intrapersonal took place in Block's Inverted Earth scenario. The story goes as follows. The Inverted Earth is just like Earth, with two exceptions: every thing has on Inverted earth the complementary color of its correspondent on Earth, and people on Inverted Earth use inversely the vocabulary about colors. When they mean red, they say green, when they say yellow, they refer to violet etc. So the sky would look orange to them but, when asked about its color, they would answer it is blue. Now, suppose a team of mad scientists put you to sleep and, while you are unconscious, they implant in your eyes color-inverting lenses and change your skin pigments. Then they take you to the Inverted Earth, where you fill a social niche exactly like the one you filled on Earth. When you wake up, admittedly, you notice no difference whatsoever. ""What it is like" for you to interact with the world and with other people does not change at all. […] As far as the qualitative aspect of your mental life is concerned, nothing is any different from the way it would had been had you stayed home"[2]. As Block points out elsewhere[3], the fact that the phenomenal character of your experiences does not modify after you wake up on Inverted Earth is one of the key features of this scenario.

These two scenarios can be used in a diverse range of purposes[4], but I think all such uses are fallacious, since these spectral inversion scenarios are incoherent. They are wrong in assuming a mathematical property of color-inversion, as I will try to show in the following.

2. What is color inversion? What do we know about it? First of all, color inversion is not the inversion of colors. Colors remain just as they are; only the perception of the observer is altered. Through a certain mechanism, he perceives instead of the "real" colors other colors, precisely connected with the real ones according to a definite rule. There are very many conceivable ways of such color alteration[5]. But what we are interested in is a color perturbance behaviorally undetectable. Thus, there remain only two plausible modes of inversion, which I will discuss next.

The first type of inversion consists in associating the "tail" of the rainbow with its "head", the second color in the spectrum with the penultimate etc. Basically, we write the spectrum backwards, we put it next to a normal spectrum, and we determine the most important correspondences that give meaning to the phrase "color inversion". I will call this kind of color-inversion "reversion".

Red – Violet;
Orange – Blue;
Yellow – Green;
Green – Yellow;
Blue – Orange;
Violet - Red.

This kind of color inversion is largely discussed in Palmer (1999), where it is argued that this inversion (the inversion of

[2] Block, "The Inverted Earth", *Philosophical Perspectives*, 4: 53-79, p. 63
[3] Review: "Is Experiencing Just Representing?", *Philosophy and Phenomenological Research* > Vol. 58, No. 3, p. 665

[4] See Alex Byrne, "Inverted Qualia", *Standford Encyclopedia of philosophy,* 2004
[5] For example, if we assign to each main color a number from 0 to 7, we can think of the following formula for color inversion function: $r(x) = x + 3$ (where 4+3=0, 5+3=1, 6+3=2, 7+3=3).

the red-green dimension of the three-dimensional color space) is empirically possible. This would mean that red-green color-inverted people are walking among us and that this color-inversion is behaviorally undetectable, because nobody ever identified such a color-inverted perceiver. The argument maintains that, given the fact there have been identified both protanopes, and deuteranopes, it is possible to conceive that some people might have both these forms of red-green color-blindness. Such people would be, thus, not color-blind, but inverted trichromats.

The other kind of color-inversion is complementarity. The inverse of a color is its complementary color. Thus, the list of fundamental correlations regarding this type of color-inversion is the following:

Red – Green;
Orange – Violet;
Yellow – Blue;
Green – Red;
Blue – Yellow;
Violet – Orange.

It is this mode of color-inversion that has been discussed the most by philosophers. It is also the most problematic. But the former too is not without problems. I think they are both vulnerable to the same kind of objection, and I will try to sow it later on.

3. Basically, when we look at the whole picture of color-inversion we notice the following elements. First of all, we have a set of different numbers, representing the values of some physical magnitude concerning the electromagnetic waves (such as the wave-length or frequency). Then we also have a set of color-classes. Those can be constructed by formal means, let's say by means of Carnap's quasi-analysis from his *Aufbau*. We have in addition three functions. One of them is the composition of the waves. Another is the transduction function, which connects numerical values corresponding to different wavelengths with color-classes. And we have a final function which "inverts" the colors.

We can easily see that the composition function has for its domain the Cartesian product of the first set with itself and for a converse domain the simple set. The transduction function connects both sets, and the inversion function operates only in the second set referred to, namely se set of color-classes. Thus, noting the first set with L and the second with C, the composition function with *c*, the transduction function with *t* and the inversion function with *r*, we can write down the following:

(1) Let *c* be a binary function such as c: L x L → L, and $c(x, y) = x + y$. We call this function "the composition operation" and we write "$x + y$" for "the wave with the wavelength x composed with the wave with the wavelength y". We don't know yet the properties of this function.
(2) Let *t* be a function, such as t: L → C. We write $t(x)$ and we read "the transduction of the light wave stimulus with the wavelength x", and we hold that the function has the following property: $\forall y \quad y \in C \rightarrow \exists x \ (x \in L \ \& \ t(x) = y)$ (surjectivity).
(3) Let *r* be a function, such as r: C → C. We call it "the inversion operation", we write $r(x)$ and we read "the inverted color-class of the transduction of the light wave with the wavelength x". The following holds for r: $r(r(x)) = x$.

We give the next definition:

D1. A *formal concept of color-inverted Earth* is a quintuple <L, C, +, *t*, *r*> which obeys the conditions stated in (1) – (3).

The formal concepts of color-inverted Earth (CIE) can be classified by the criterion of the inversion operation. As we have previously seen, there are four conceivable inversion functions. It follows that we have two different formal concepts of CIE, and these are:

Concept 1. <L, C, +, *t*, *reversion*>
Concept 2. <L, C, +, *t*, *complementarity*>

What is missing from the picture is in both cases the set of logical properties for the composition operation. And also, we need to know something about the physical and mathematical meaning of the inversion functions. I will investigate these issues closer in the next section of this essay.

4. The principles of color composition were first investigated in Newton's *Optics*. The famous representation of the color wheel[6] offers some suggestions for color composition and, indirectly, color inversion.

As it is well known, Newton arranged the colors of the rainbow on a circle, and divided the circle in 7 parts, according to a musical analogy. Actually, there is no good reason to include indigo among the principal colors of the spectrum. However, Newton obtained seven arcs, each representing the range of a color, in the proportions of the seven musical intervals corresponding to the eight sounds. In order to compose colors, Newton developed an intuitive method in his VI-th proposition, Book I, Part II of his *Optics*. He said about it: "This rule I conceive accurately enough for practice, though not mathematically accurate"[7].

The rule of composition needed to know the quantity and quality of the two colors to be mixed, and could calculate the color to be obtained. Newton constructed each arc's gravity center, and combined colors by calculating in conformity with the two colors' quantity the center of gravity of the segment uniting the centers of gravity of the two arcs representing colors to be combined, and then tracing a radius of the circle through this newly obtained point, representing the center of gravity of the segment. The point of intersection of the radius thus traced with the circle's circumference indicated the color obtained by this composition. This rule proved extremely accurate empirically; this has been proved by J. Clerk Maxwell, trough a series of experiments communicated to the *Royal Society of London*[8].

What Newton first discovered about colors was that there were two kinds of color properties: optical and chromatic. Accordingly, there were two kinds of color: monochrome (or homogenous) and composed. Their chromatic properties are not criteria of individuation; practically, we have no possibility of distinguishing between a monochrome and a compound green on the basis of their chromatic properties. But we can discern between them, on the basis of their optical properties, by using a prism. The compound green will split through the prism in blue and yellow, while the monochrome red will not be affected by going through the prism. This distinction is very important, since color composition is mainly a chromatic operation.

Grassman was the next to bring an important innovation in the theory of compound colors, by showing that colors combine like vectors, according to the rule of parallelogram[9].

Maxwell, in change, developed an algebraic treatment of color composition. He showed how we can specify the relation between any four colors, by the following equations:

(1) $u=x+y+z$;
(2) $u+x=y+z$;
(3) $u+x+y=z$.

These equations say that any color either can be obtained from three colors, either can be mixed with another color to obtain a new color, obtainable also through a different mixture of colors. For example, red combined with blue gives the same color as orange combined with green. But these relations are not explanatory; they merely

[6] Isaac Newton, *Opticks*, London: Sam. Smith & Benj. Walford, MDCCIV, p. 114 *sq.*
[7] *ibidem*, p. 117

[8] J. Clerk Maxwell, "On the Theory of Compound Colors, and the Relations of the colors of the Spectrum", *Transactions of the Royal Society of London*, vol. 150 (1860), pp. 57-84
[9] Maxwell, *op. cit.*, p. 60

capture the relations found in Newton's diagram, in another way. But we are far from a complete explanation of Newton's great intuitions.

The interesting thing about Maxwell's relations is that they can express what it means that two colors are complementary. Given the fact that, in order to consider two colors complementary, their mixture must yield white, we can define x's complementary by the formula $x+y = white$.

Another way of expressing complementarity is to formulate the trigonometric relations between the points corresponding on the circle to the two colors. Again Grassman noticed that in the visible spectrum there must be an infinite number of pairs of complementary colors. That means that each diameter of the circle determines a pair of complementary colors. Then it follows that the relation between complementary colors can be represented mathematically by the following formula: $\sin x - \sin y = 1$, where x and y are measures of angles corresponding to the two colors' projection on the circle.

I suppose there could be developed other methods of expressing the relation between two complementary colors. But what is wrong with all these methods is that they just describe the properties of Newton's diagram, and nothing else.

Color composition is a chromatic operation, which is why physics can't tell us anything about it. The composition of waves is something; the composition of colors, something else. There is nothing we can deduce about color composition from the physical relations describing the composition of electromagnetic waves, since it is nothing intrinsic that makes a wave have chromatic properties. Similarly, there is nothing intrinsic in an electromagnetic wave, which makes the color corresponding to it be the complementary of other color. Sure, the chromatic relation between two complementary colors is known: their mixture results in white. But the radiation corresponding to white itself is by nothing intrinsic privileged among other radiations of the electromagnetic spectrum. So, much of what we know about color composition has no physical interpretation, and no physical explanation. However, since our color space is structured (its structure being determined by the structure of our perceptual apparatus), there is a certain logic in color composition. But what logic is there in color inversion? Or rather, what mathematic is there?

Our concept of color-inverted earth gets rid of the physical and psychological problems of the color inversion[10]. What remains is purely the problem of making sense of our concept of color complementarity in mathematical terms. It is the mathematic of color complementarity that eludes us. So, how can we do that? It is not hard to notice that color composition can be treated in terms of set intersections. Since our color-classes are fuzzy classes, color intersection will be fuzzy intersection. But what can be said about the inversion of the colors?

First of all, if we take inversion to be what I called "reversion", we can say nothing at all. We can not give mathematical meaning to a function that empirically maps red to violet, orange to blue and so on. That is, we can't give the function through a general formula, but only through a pair list.

Second, its looks like there is more sense in treating inversion as color complementarity. Complementarity seems to be a more rigorous notion. But is there really more to complementarity than another pair list? Every color-class has a complementary class. But, green being our color-class, its complementary class is the class of non-green, that is, the reunion of all other color-classes. So clearly, the complementary class of a color-class is not the same thing as the complementary-color class we look for. But the complementary-color class is a subclass of

[10] The transduction function is like a black box, thus eliminating any concern with problems of perception. By replacing colors with color-classes we again escape the need of paying attention to psychology or neurology.

the complementary color-class. But the reversed-color class had exactly the same status, i.e. of being a subclass of the complementary color-class.

So, is there some mathematical property which distinguishes the complementary-color class of a color-class by its reverted-color class? This question reduces to the following: does the complementary-color class of green have an *a priori* relation with the green color-class? Again, we have at our disposal the following answer: Yes, it does. We know *a priori* that green color-class intersected with its complementary-color class makes the white color-class (or, which is the same, that green mixed with its complementary color gives white). This would seem to move the problem from the relation between a color-class and its complementary-color class to its closure: the relation between a color-class and the white color-class. And this is a relation about which we know nothing *a priori*, except the basic fact of one's being included in the other.

Thus, I think there is no possibility of giving physical or mathematical meaning to color inversion, however we look at what inversion is. All we have is just some pair lists, empirically obtained. And this, I think, has important consequences on the possibility of the inverted spectrum, that is, on the validity of the formal concepts of color-inverted earth.

5. Let's turn back to our two concepts mentioned in the section 3. We have to enforce on them the condition of validity, strongly connected with the impossibility of behavioral detection of color inversion. I will say that a formal concept of color-inverted Earth is valid if and only if 1) is not contradictory and 2) the parallelism between the composition of colors and the composition of their inverse colors can never be broken. That is exactly what "impossibility of behavioral detection of inversion" means. If we can produce a counter-example which breaks this parallelism, given what inversion is in each concept, then we can prove that the formal concept of color-inverted Earth is not valid. And this counter-example can be easily produced.

Suppose Inverted Earth is possible. Then we must accept by hypothesis the following statements:

(1) On Earth the person to travel to Inverted Earth perceived the color normally.
(2) On Inverted Earth, due to the lens that annulates the color inversion, *he will not perceive any difference in his colored experiences*.

Now suppose the emigrant was, on Earth, a painter in aquarelle. He used to prepare his own colors, by mixing the six principal colors of the spectrum. He procured his colors under the form of some pigment disks, from a self-servicing store. The night before he was kidnapped, our painter had to fill a surface of a painting with violet, color he prepared by mixing in equal proportions red and blue. Let's suppose further that, because of his tiredness, the painter goes to bed after painting with violet only a half of the surface intended.

At night, the famous team of mad scientists implant in his eyes color-inverting lenses, makes a copy of his painting in complementary colors and transport him along with the copy of the painting on Inverted Earth, where they substitute him to his doppelganger.

The next morning, our painter returns to his painting. He prepares as usual his violet, mixing what he perceives as red with what he perceives as blue. Due to his being on Inverted Earth, and not on Earth, what he is really mixing is, in case of *Concept 1*, violet and orange, and, in case of *Concept 2*, green and yellow. The violet from earth was replicated on the copy painting with its inverse color, namely red, in a case, yellow in the other. But what the painter obtains and spreads with its paintbrush near this red/yellow, is either reddish-brown, or yellowish-brown. And those are NOT the same thing either with red, or with yellow.

At this stage of the experiment there can happen, at first sight, two things: either the painter suffers a shock seeing how *his combination of red and blue gives something different from violet*, or the painter will notice absolutely no difference between the two colors on the canvas. The first alternative would break the parallelism between the two compositions, thus invalidating both concepts. But the breaking of the parallelism is forbidden from hypothesis of the experiment, by the claim (2). So only the second alternative is allowed.

Similarly, for *Concept 2*[11], by mixing red and orange to obtain reddish-orange, the painter will actually mix green and blue and he will get cyan, which is different by the complementary of reddish-orange – something between blue and violet. But in this case too, he will not be able to notice the difference.

This means that from the painter's multitude of color experiences a series of experiences have disappeared, without the painter's being able to notice. That means the simple journey on Inverted Earth has transformed him in a partial zombie, for the shrinking of his field of colored experiences in significant. By going further with our imagination we can catch a glimpse of how, because the painter is suddenly incapable to discern yellow from yellowish-brown, he might combine yellowish brown with blue, thinking he mixes violet with red in order to obtain indigo. In reality he will get some kind of greenish-brown he will not be able to discern from reddish-orange, the complementary of the indigo. Step by step, if we follow all the possible combinations of colors, as the painter perceives them, we reach the following reasoning:

(i) the painter perceives (by hypothesis) all the color perceivable by a normal eye, even if inverted;

(ii) the hypothesis implies, as showed, that the painter can only perceive the principal colors of the spectrum;

(iii) moreover, it is plausible the painter can not perceive even all principal colors of the spectrum.

(iv) as it is noticeable, (ii) and (iii) contradict (i).

(v) That means the hypothesis of the Inverted Earth implies affirmations that contradict it, which shows that there is no valid concept of color-inverted Earth.

6. This argument exhibits a mathematical property of the inversion function. In usual language, we can express that by saying that the mixture of complementary colors is not always identical with the complementary of their mixtures. That means, complementarity (or, largely, inversion) is not distributive to composition. Thus, what the argument shows is the truth of the following relation:

(R) $\exists x, y\,[\mathrm{comp}(x \cap y) \neq (\mathrm{comp}\,(x) \cap \mathrm{comp}\,(y))]$

Unfortunately, there is no way of proving the necessity of (R)[12]. Therefore the argument presented fails to support the metaphysical impossibility of this scenario. It only shows that, based on our color experience, the mathematics of complementarity doesn't allow us to coherently conceive such a scenario. But it leaves open the possibility that some other creatures, with different visual sense organs, with different perceptions, infringe (R). Those creatures might not perceive any difference in their color experiences just by traveling from their planet to some inverted planet. However that may be, the argument is still useful

[11] I will focus from now on, for simplicity, only on the *Concept 2*.

[12] Since all we know about color-complementarity is empirically grounded.

against diverse criticisms brought to functionalism or representationalism about qualia.

References:

Block, Ned (1990) - "Inverted Earth", *Philosophical Perspectives*, 4: 53-79.
Block, Ned (1998), Review: "Is Experiencing Just Representing?", *Philosophy and Phenomenological Research* 58, No. 3: 663-670
Byrne, Alex (2004) – "Inverted Qualia", *Standford Encyclopedia of Philosophy*, http://plato.stanford.edu
Chalmers, Davis (1996) - *The Conscious Mind*, Oxford: Oxford University Press
Dennett, Daniel (1991) - *Consciousness Explained*, Boston: Little, Brown
Maxwell, Clerk J. (1860) - "On the Theory of Compound Colors, and the Relations of the colors of the Spectrum", *Transactions of the Royal Society of London*, 150: 57-84
Newton, Isaac (1704) – *Opticks*, London: Sam. Smith & Benj. Walford
Palmer, Stephen (1999) - "Color, Consciousness, and the Isomorphism Constraint", *Behavioral and Brain Sciences*, 22: 923-43.
Shoemaker, Sydney (1982) - "The Inverted Spectrum", *Journal of Philosophy,* 79: 357-81