

$$\begin{aligned}
N_{12} &= \sum_{j=1}^{\infty} N_{12}^j \\
&= N_{12}^1 + \sum_{j=2}^{\infty} N_{12}^j \\
&= N_{12}^1 + \sum_{j=2}^{\infty} (N_{21}^{j-1} \Sigma_{212}) \\
&= N_{12}^1 + \left(\sum_{j=1}^{\infty} N_{21}^j \right) \Sigma_{212} \\
&= N_1^0 \Sigma_{12}^0 + N_{21} \Sigma_{212} \tag{13}
\end{aligned}$$

The last equality follows from (12) and the preceding definition of N_{12} . By comparing (13) and (4) we see that the natural mode of solution implies the interaction mode of solution. Evidently the steps in (13) are reversible, so that the interaction mode of solution implies the natural mode of solution. Thus the two modes of solution are equivalent in this case. Since the interaction mode of solution clearly represents the solution of the interreflection problem of S_1 and S_2 , the natural mode of solution therefore is also, by virtue of the preceding equivalence, a solution of the interreflection problem. This equivalence actually holds in very general settings and has been established in detail for these settings, in Ref. [251]. We shall have occasion to study and use once again this equivalence of the two techniques later in the present work. Finally, we observe that the sums in (11) and (12), being reducible to a simple geometric series with ratio $\Sigma_{121}\Sigma_{212}$ and initial term of the form $(N_1^0 \Sigma_{12}^0 + N_j^0 \Sigma_{ji}^0 \Sigma_{jij})$ ($i=1, j=2$ for (11); $i=2, j=1$ for (12)), are readily evaluated; these sums are given by (6) and (7).

3.2 The Interaction Principle

With the preliminary example complete, we turn now to the statement of the central principle of radiative transfer theory:

The Interaction Principle: For every X, S, A, B, m and n , if X is an optical medium and S is a subset of X , and A ($= (A_1, \dots, A_m)$) is a class of sets A_i consisting of incident radiometric functions on S , and B ($= (B_1, \dots, B_n)$) is a class of sets B_j consisting of response radiometric functions on S , and m and n are positive integers, then there exists a unique set $\{s_{ij} : i=1, \dots, m, j=1, \dots, n\}$ of linear (interaction)

operators s_{ij} with domain A_i and range B_j with the property that for every element (a_1, \dots, a_m) of A there exists an element (b_1, \dots, b_n) of B such that:

$$b_j = \sum_{i=1}^m a_i s_{ij}$$

or in matrix form:

$$b = as$$

where we have written:

"a" for (a_1, \dots, a_m)

"b" for (b_1, \dots, b_n)

"s" for
$$\begin{pmatrix} s_{11} & s_{12} & \dots & s_{1n} \\ s_{21} & s_{22} & \dots & s_{2n} \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ s_{m1} & s_{m2} & \dots & s_{mn} \end{pmatrix}$$

Discussion of the Interaction Principle

We shall discuss in some detail the meanings of the various terms in the interaction principle. First of all, the meaning of the term "optical medium" as used in the statement is quite broad and, for example, is intended to have as real designata such parts of the world as lakes, oceans and various portions of the atmosphere. From the mathematical point of view, "optical medium" may be interpreted simply as part of Euclidean three-dimensional space such as the region between two infinite parallel planes or the interior of a sphere, etc., in which we assume that the principles of geometric optics hold, in particular, Fermat's principle. There will eventually evolve, as the studies progress and the basic constructs assume their final form, a relatively technical version of what we mean by the term "optical medium" in the fully developed theory (re: Def. 5 of Sec. 9.1). However, for the present the term may have either of the simple meanings suggested above.

The meanings of the terms A , B , and s_{ij} in the principle can be illustrated using the preliminary example of Sec. 3.1. Let us return to the setting summarized by Eqs. (4) and (5) of Sec. 3.1. In that setting the optical medium was some (physically) vacuous region X of Euclidean three-space containing two plane surfaces S_1 and S_2 . We concentrate attention on S_1 . Then S_1 is an instance of S in the principle. Consider the set of all incident radiances like N_1^0 on S_1 .

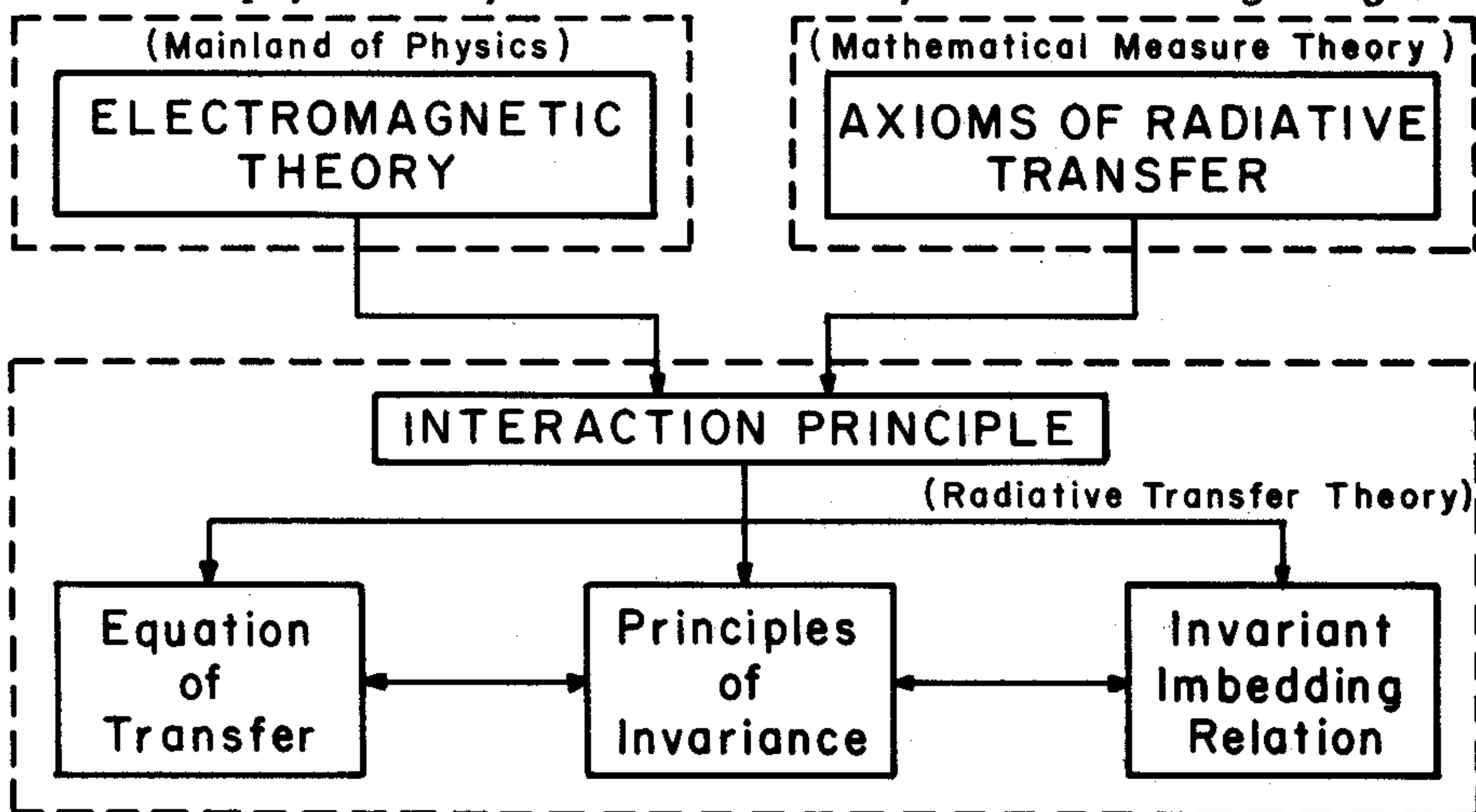
This set of incident radiances becomes the set A_1 in the principle. Consider the set of all incident radiances like N_{21} on S_1 . This becomes the set A_2 in the principle. Together, (A_1, A_2) constitute the incident class A in the principle, so that $m = 2$. It should be noted that A_1 and A_2 are each closed under the operations of forming sums and products by nonnegative numbers (*linear closure*). Thus if N_1 and N_2 are in A_1 , then so is $cN_1 + dN_2$ where c and d are nonnegative numbers. This feature of A_1 and A_2 comes automatically with the requisite linearity of the s_{ij} . The class B of response functions S_1 consists of one set B_1 , with N_{12} as a typical element. Therefore in the case of S_1 we have $m = 2$, and $n = 1$, with Σ_{21}^0 and Σ_{212} as the present instances of s_{11} and s_{21} , respectively. Hence one invocation of the interaction principle for the case of S_1 yields (4). Another and distinct invocation in the case of S_2 yields (5).

The alternate example summarized in (10) of Sec. 3.1 provides a further illustration of the principle's linear algebraic statement. In (10) of Sec. 3.1, X is the same space as above. Now, however, S_1 and S_2 are considered parts of one and the same subset, say S of X . Consider the set of all ordered pairs of incident radiance on S like (N_1^0, N_2^0) . This becomes A_1 in the principle. Consider the set of all ordered pairs of response radiances of S like $(N_{1\beta}, N_{2\gamma})$. This becomes B_1 in the principle. Therefore in the present case of S , we have $m = n = 1$, and s_{11} is $\phi_{\beta\gamma}^0$. As we select any new incident pair (N_1^0, N_2^0) , there corresponds the associated response pair $(N_{1\beta}, N_{2\gamma})$ given by (10). Clearly (10) is the present instance of the matricial form of the principle's algebraic statement.

As we progress along the line of examples of the interaction principle we shall be gradually less explicit in pointing out the particular parts of the current form of the interaction principle, leaving the details of correlation more to the reader as he becomes familiar with the principle. In all the subsequent uses of the principle, we shall look upon it as a convenient working principle, i.e., a rule of action for the formulation of subordinate principles, the various laws, and everyday problems of radiative transfer theory. The practical uses of the principle are directed to determining the light field in natural optical media by finding the interaction operator s_{ij} , supplied by the basic principle, for a given medium. The determination of the structure of the operators s_{ij} and the various functional equations they satisfy constitutes one of the more interesting and challenging problems of modern radiative transfer theory. We shall begin the investigation of these operators in the present chapter and continue it in Chapter 7.

The Place of the Interaction Principle in Radiative Transfer Theory

It is not intended that the interaction principle categorically replace all classical instances of itself such as the principles of invariance and the invariant imbedding relation, or other classical instances that occur in the literature or that arise during the subsequent developments below. Rather, it is intended that the principle be viewed by its users simply as a working principle of radiative transfer theory, and to be used (and perhaps refined) by those students of the subject who prefer to envision the theory as governed by and derivable from a single idea. The place of the interaction principle in radiative transfer theory and in the mainstream of physics may be summarized by the following diagram:



As the diagram indicates, radiative transfer theory may join the mainland of physics via electromagnetic theory (see, e.g., Chapter XIV, Ref. [251]) or the theory may be made completely autonomous using an axiomatic formulation made elsewhere (Chapter XV, Ref. [251]). Direct interconnections also exist between the three principal parts of the theory (indicated in the diagram below the interaction principle). In fact the internal ties on the level of the general equation of transfer, the general principles of invariance, and the general invariant imbedding relation are so strong that these ties are effectively logical equivalences. The details of the pursuit of these connections are mainly mathematical and are beyond the scope of the present work. For further details on this matter, the reader is referred to the various chapters of Ref. [251].

Levels of Interpretation of the Interaction Principle

The great practical range and depth of the interaction principle arises from the levels of interpretation on which it may be applied. There are generally four main levels of

interpretation of the principle: the *point*, *line*, *surface*, and *space* levels. Of these, the surface and space levels of interpretation are operationally the most meaningful. The point and line interpretations are special theoretical artifices which increase the range of the principle in specific settings. The preliminary example above is an instance of the surface level of interpretation. In general, the *surface level* interpretation of the interaction principle subsists when one interprets the subset S of a space X as a subset of one less dimension than X . For three-dimensional spaces X , S would have two dimensions. For two-dimensional spaces X (which arise in certain mathematical models) S would have one dimension, etc. In general the *space-level interpretation* of the interaction principle subsists when one interprets the subset S of a space X as a subset of the same dimension as X . Plane-parallel slabs, spherical solids in Euclidean three space are settings for the space-level interpretation. For two-dimensional spaces X , the subset S would have two dimensions, etc.

Of the remaining two levels of interpretation of the principle, the point level interpretation is the more widely used. In fact the point-level interpretation covers so much ground that it is convenient to regard it from two separate aspects. The *general point-level interpretation* of the interaction principle subsists when X is a general space whose points are arbitrary. The general point-level interpretation is of most use in the development of general discrete-space theory (Ref. [251]). The *special point-level interpretation* of the interaction principle subsists when S is a point or an optically small three-dimensional subset of space (i.e., e.g., a point source) in which single scattering processes are to be dominant relative to multiple scattering processes. This special interpretation is commonly used to establish in an intuitive fashion the concept of the volume scattering function, which plays a key role in the theory (see Sec. 13.4). An alternate establishment of the volume scattering function could take place strictly and rigorously in the space-level interpretation (see Sec. 3.14). The special point-level interpretation is also a useful and defensible ploy in setting up radiative transfer theory and is thereby retained and given a special status. (See, e.g., Example 1, Sec. 3.17.)

The final level of interpretation to be discussed is the line-level interpretation of the interaction principle. The *line-level interpretation* subsists when one interprets the subset S of a space X as a one-dimensional subset of X . The line-level interpretation is not operationally meaningful as are the surface, space and special point-level interpretations. However, it is retained because it favors useful mathematical artifacts, as does the special point-level interpretation. Furthermore, like the special point-level interpretation, the use of the line-level interpretation is rigorously defensible by means of limit arguments starting with the space-level interpretation; for that reason it is retained as a useful technical device. We shall use it below in viewing the path radiance as the response of a path in real optical medium to the incident path function radiances along the path. (Example 2, Sec. 3.17.)

Unless specifically noted otherwise, we shall henceforth mean by "optical medium" any three-dimensional part X of Euclidean three-dimensional space. This then will automatically set the dimensionality of S in the various interpretations of the interaction principle. (A formal definition of optical media, as they are studied in radiative transfer theory, is given in Sec. 9.1.)

3.3 Reflectance and Transmittance Operators for Surfaces

In this section we begin the sequence of constructions of the concepts needed for the description of the manifold radiative transfer phenomena encountered in the practice of radiative transfer theory. In particular in this section we shall use the interaction principle as a base for the construction of the more commonly used surface reflectance and transmittance concepts. Some work has already been done in this direction in Sec. 3.1. In fact the empirical reflectance function was defined in that section as a necessary prerequisite for the construction of the preliminary example of the interaction principle. We now return to that setting for the purpose of establishing systematic definitions for the family of reflectance and transmittance operators for surfaces.

Geometrical Conventions

Figure 3.3 (a) depicts a general surface Y in an optical medium X and a relatively small part S of Y about point x on Y . We are interested in the reflectance and transmittance of Y in the region S about x . Now the terms "transmittance" and "reflectance" become meaningful only after adequate reference frames have been established at given points x of Y within which one can unambiguously establish conventions about the notions of "inwardness", "outwardness", "upwardness", "downwardness", "forwardness", "backwardness", etc. Suppose then we affix to point x of Y a unit vector $k(x)$ and call it the *unit outward normal* to Y at x . Perhaps some readers would prefer to call $-k(x)$ the unit outward normal to Y at x . This is perfectly admissible for our present purposes, and the reader may therefore turn around the arrows in parts (a)-(d) of Fig. 3.3 and read the following discussion as it stands. The point being made here is that what one calls "outward", etc., is immaterial. What does matter is what one subsequently does with the concept and that, within a given discussion, a measure of consistency is sustained in the use of the concept once the convention is made.

During the present discussion, let " D' " and " D " denote narrow circular conical solid angles of central directions ξ' and ξ , respectively. S is a small collecting surface on Y , and x is a point of Y in S . Let " S' " denote the projection of S on a plane normal to ξ' . (See parts (c) and (d) of Fig. 3.3.) D' is the set of *incident directions*; D is the set of *response directions*. Both D' and D will always lie completely within $E_+(k(x))$ or $E_-(k(x))$ where $E_+(k(x))$ is the set of all directions ξ' such that $\xi' \cdot k(x) > 0$, and $E_-(k(x))$ is the set