Princeton Plasma Physics Laboratory

Optimal Shielding Design for Minimum Materials Cost or Mass

Robert D. Woolley

July 2015

PPPL
PRINCETON
PLASMA PHYSICS
LABORATORY

Princeton Plasma Physics Laboratory
Report Disclaimers

Full Legal Disclaimer

Trademark Disclaimer

PPPL Report Availability

Princeton Plasma Physics Laboratory:

http://www.pppl.gov/techreports.cfm

Office of Scientific and Technical Information (OSTI):

http://www.osti.gov/scitech/

Related Links:

U.S. Department of Energy

U.S. Department of Energy Office of Science

U.S. Department of Energy Office of Fusion Energy Sciences

Optimal Shielding Design for Minimum Materials Cost or Mass

Robert D. Woolley

Princeton Plasma Physics Laboratory, Princeton University, POB 451, Princeton, NJ 08543

woolley@pppl.gov

Abstract

The mathematical underpinnings of cost optimal radiation shielding designs based on an extension of optimal control theory are presented, a heuristic algorithm to iteratively solve the resulting optimal design equations is suggested, and computational results for a simple test case are discussed.

A typical radiation shielding design problem can have infinitely many solutions, all satisfying the problem's specified set of radiation attenuation requirements. Each such design has its own total materials cost. For a design to be optimal, no admissible change in its deployment of shielding materials can result in a lower cost. This applies in particular to very small changes, which can be restated using the calculus of variations as the Euler-Lagrange equations. The associated Hamiltonian function and application of Pontryagin's theorem lead to conditions for a shield to be optimal.

Keywords: shielding, optimization, Pontryagin

I. INTRODUCTION

Materials costs dominate some shielding design problems. This is the case for manned nuclear power space applications for which shielding is essential, for example, and the cost of launching by rocket from Earth is high. In such situations or in situations where shielding volume or mass is constrained, it is important to optimize the design. Although trial-and-error synthesis methods may succeed, a more systematic approach is warranted. For other applications with lower effective shield materials cost rates, optimal design involves an automated approach that may reduce design labor costs.

Traditional approaches to shielding design focused on achieving adequate attenuation. Early in the nuclear age, computer prediction of a shield's attenuation was expensive and inaccurate, so direct measurement was frequently relied on. Attenuation by a pure absorber is easy to accurately calculate, but scattering is more complicated. Approximate methods were developed such as the use of buildup factors summarizing scattering results of accurate attenuation calculations or measurements. Buildup factors allow rough prediction of shield attenuation using simplified calculations without detailed consideration of the radiation's energy spectrum. Removal cross sections were another approximation developed for situations in which fast neutrons are first scattered while passing through a material slab before entering a thick water shield. They allowed the slab's scattering effects on subsequent water absorption to be modeled as an additional pure absorption without detailed consideration of the neutron energy spectrum.

A 1973 compendium[1] detailed the shielding design methods then in use. Its brief discussion of shielding design optimization (on 3 of its 788 pages) reports five approaches of which three

implemented algorithms derived from a variational approach and a fourth implemented a gradient nonlinear programming scheme.

The fifth approach was the ASOP computer code,[2] new at that time, which minimized mass of spherical shields for a given attenuation requirement through a direct numerical search. ASOP evaluated each shield configuration's attenuation by running the newly developed ANISN discrete ordinates code,[3] then evaluating finite differences between successive ANISN runs using slightly varied radial boundaries to guide its next shield redesign iteration, and continuing iterations until ASOP's convergence criteria were met. However, ASOP did not fully automate shield design since it required a wise selection of shielding materials sequenced into layers. Fifteen years later, shielding optimization techniques had not progressed. A shielding optimization project for satellites documented in a 1988 report[4] focused on using the same ASOP code to optimize selected combinations of project-selected shielding materials into spherical layers. No fully automated approach to optimization was used. A 1996 textbook[5] on shielding continued to ignore optimization while focusing on shield attenuation calculation methods.

By now, the accuracy of shielding computations has enormously improved with better numerical methods, cross section libraries and bench-marked standardized codes, while computation costs have enormously decreased. However, shield optimization methods have not evolved. The present work is therefore intended to spur improvements in shield design methods to include shield optimization. The eventual objective is to fully automate design of optimal radiation shielding in any geometric configuration to meet specified attenuation requirements given a menu of allowed shielding materials along with their properties and cost densities.

As developed herein, the cost-optimal design turns out to be the one which, at each point location in the shield, chooses a material minimizing a particular functional of forward and

adjoint angular fluxes plus material cost density. Thus a Boltzmann Transport Equation (BTE) forward and adjoint angular flux solver augmented with some additional calculations as described herein can evaluate whether a shield design is optimal or if improvements are possible. An iterative algorithm based on approximated optimality is proposed to solve for the optimal shield design, and numerical experiments with it are discussed.

I.A. Optimal Control Theory Background

The starting point for optimal shielding design is optimal control theory, a field with an extensive literature that developed from control theory. Control theory had emerged within electrical engineering during the first half of the last century to regulate continuous processes via analog electronics using circuit integrators, amplifiers and sensor feedback. Controlled processes were typically linear time-invariant, so frequency-domain spectral methods from radio engineering had ensured that servo loops responded without oscillation.

Coincident with rocket-powered missile development, control theory expanded to include state-space methods. The state of a system is an ordered set of scalar variables with the property that the state at any instant allows prediction of subsequent responses to control inputs or disturbances without knowing earlier history. In continuous-time systems the state variables change smoothly without discontinuous jumps in contrast with input control commands which may vary arbitrarily.

As a simple example, the state of a swinging pendulum with applied horizontal control force is described by two state variables, angular position and angular rate. A more difficult example considers controlling the internal temperature of an object by heating its boundary, for which the precise state of the system at any instant requires representing temperatures at all internal points. This need for infinite state data in order to make exact response predictions is typical for

distributed parameter systems described by partial differential equations (PDEs). For engineering purposes such systems are usually represented approximately using a finite number of state variables, e.g., modeling internal temperatures at only a finite number of spaced grid points.

In the practical example of a rocket-powered missile during its controlled flight, the state variables include its remaining fuel mass and its three components each of position, velocity, rotational orientation and angular velocity. However, these thirteen state variables are joined by others modeling internal subsystems controlling, e.g., aerodynamic control fins, a rocket engine gimbal and/or a throttle valve.

In state-space analysis of systems with a finite number of state variables, the system's physics are incorporated as a set of first-order ordinary differential equations (ODEs) expressing the time rate of change of each state variable as a function of all state and control variables, possibly with additional dependencies on time. It is conventional to refer to a system's n state variables as its n-dimensional state vector, $\underline{x}$, to its m control variables as the m-dimensional control vector, $\underline{u}$, and to its n rate-of-change functions as the n-dimensional vector function, $\underline{f}$. Per Eq. (1), the time-varying control vector's range is constrained to a specified set, U, whose boundaries if any represent inherent technical limits:

$$\underline{u}(t) \in U \subseteq \Re^m \quad . \tag{1}$$

The system dynamics are expressed as the following vectorized state equation:

$$\frac{d}{dt}\underline{x}(t) \equiv \underline{\dot{x}}(t) = \underline{f}(\underline{x}(t), \underline{u}(t), t) \quad . \tag{2}$$

Eq. (2) is accompanied by a boundary condition such as the initial state condition:

$$\underline{x}(t_0) = \underline{x}_0 \quad . \tag{3}$$

The system's response to a control input is a trajectory through n-dimensional state-space. When evaluating possible control strategies starting from a specified initial state, only admissible pairs of state and control trajectories $(\underline{x}(t), \underline{u}(t))$ satisfying the physical constraints of Eqs. (1) and (2) are considered.

Adoption of state-space analysis methods brought a new emphasis on optimization, e.g., for maximizing the altitude or speed reached by a rocket, or for minimizing a rocket's use of fuel or energy or time needed to reach a target or nearest approach distance to a missed target. These examples of different quantities to optimize can each be expressed as the sum of a specified scalar performance function of the final state reached, $\varphi(\underline{x}(t_f), t_f)$, plus an integral over the mission time interval of some other specified scalar performance rate function of the path taken, $L(\underline{x}(t), \underline{u}(t), t)$. Thus, the general optimal control problem is to choose $\underline{u}(t)$ to maximize (or minimize) the following scalar performance index functional of the entire histories of the state and control functions:

$$ J[\underline{x}(t); \underline{u}(t)] = \varphi(\underline{x}(t_f), t_f) + \int_{t_0}^{t_f} L(\underline{x}(t), \underline{u}(t), t) dt \text{ ,} \tag{4} $$

subject to the state equation Eq. (2) boundary conditions including Eq. (3) and problem-specific constraints including Eq. (1).

To optimize a mission its time interval must be well defined, but it usually begins as in Eq. (3) with specified initial time t_0 and state $\underline{x}_0$. However, a wide variety of situations are treated for the final time, with the final time itself in some cases specified in advance, in other cases determined as the time when specified conditions are reached, and in yet other cases left as a free parameter to be optimized. It turns out that a general framework approach can be taken and the

different boundary condition details of different situations affect only corresponding boundary condition details of their optimal solutions.

The optimal control solution is pursued by affixing Eq. (2) to the Eq. (4) functional's integrand using the Lagrange multiplier $\underline{\lambda}(t)$, an undetermined n-dimensional vector function of time, as follows:

$$J = \varphi\left(\underline{x}(t_f), t_f\right) + \int_{t_0}^{t_f} \left[L(\underline{x}(t), \underline{u}(t), t) + \underline{\lambda}^T(t)\left\{\underline{f}(\underline{x}, \underline{u}, t) - \underline{\dot{x}}(t)\right\}\right] dt . \tag{5}$$

For any admissible pair $(\underline{x}(t), \underline{u}(t))$ satisfying Eqs. (1), (2) and (3), the Lagrange multiplier $\underline{\lambda}(t)$ may be chosen arbitrarily since within Eq. (5) it is multiplied by zero. Eq. (5) is next integrated by parts, and the Hamiltonian function is identified in the result:

$$J = \varphi\left(\underline{x}(t_f), t_f\right) - \underline{\lambda}^T(t_f)\underline{x}(t_f) + \underline{\lambda}^T(t_0)\underline{x}(t_0)$$

$$+ \int_{t_0}^{t_f} \left[\underbrace{L(\underline{x}(t), \underline{u}(t), t) + \underline{\lambda}^T(t)\underline{f}(\underline{x}, \underline{u}, t)}_{\substack{H(\underline{x}(t), \underline{u}(t), \lambda(t), t) \\ \text{Hamiltonian}}} + \underline{\dot{\lambda}}^T(t)\underline{x}(t) \right] dt . \tag{6}$$

Given a particular admissible state and control pair $(\underline{x}(t), \underline{u}(t))$ that satisfies Eqs. (1), (2), and (3), optimality of this pair is equivalent to the impossibility of further improving the performance functional of Eq. (5) by using any different admissible pair $(\underline{x}(t) + \underline{\delta x}(t), \underline{u}(t) + \underline{\delta u}(t))$, where $(\underline{\delta x}(t), \underline{\delta u}(t))$ represents a combination of variations in state and control histories allowed by problem constraints. For a fixed final time this optimality condition is stated in Eq. (7).

$$\delta J \le 0 \qquad \text{for a maximizing optimum} \qquad\qquad\qquad (a)$$

$$\delta J \ge 0 \qquad \text{for a minimizing optimum} \qquad\qquad\qquad (b)$$

where

$$\delta J \equiv J\big[\underline{x}(t)+\underline{\delta x}(t);\underline{u}(t)+\underline{\delta u}(t)\big]-J\big[\underline{x}(t);\underline{u}(t)\big]$$

$$= \varphi\big(\underline{x}(t_f)+\underline{\delta x}(t_f),t_f\big)-\varphi\big(\underline{x}(t_f),t_f\big)-\underline{\lambda}^T(t_f)\underline{\delta x}(t_f)+\underline{\lambda}^T(t_0)\underline{\delta x}(t_0)$$

$$+\int_{t_0}^{t_f}\begin{bmatrix} H\big(\underline{x}(t)+\underline{\delta x}(t),\underline{u}(t)+\underline{\delta u}(t),\lambda(t),t\big)\\ -H\big(\underline{x}(t),\underline{u}(t),\lambda(t),t\big)+\underline{\dot\lambda}^T(t)\underline{\delta x}(t)\end{bmatrix}dt \qquad (c)$$

$$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad . \qquad (7)$$

Since for an optimal control, Eq. (7) holds for all admissible pair variations, it in particular holds for variations of small amplitude. For sufficiently small amplitude variations Eq. (7c) can be restated using partial derivatives:

$$\delta J = \left[\left(\frac{\partial\varphi}{\partial\underline{x}}-\underline{\lambda}^T\right)\delta\underline{x}(t)\right]_{t=t_f} + \left[\underline{\lambda}^T(t)\delta\underline{x}\right]_{t=t_0} + \int_{t_0}^{t_f}\left[\left(\frac{\partial H}{\partial\underline{x}}+\underline{\dot\lambda}^T\right)\delta\underline{x}(t)+\frac{\partial H}{\partial\underline{u}}\delta\underline{u}(t)\right]dt . \qquad (8)$$

However, $\underline{\delta x}(t)$ and $\underline{\delta u}(t)$ are not independent since they are related through the Eq. (2) constraints. It would be tedious to disentangle their mutual dependencies so instead one simply chooses the arbitrary Lagrange multiplier vector $\underline{\lambda}(t)$ to make the Eq. (8) $\delta\underline{x}(t)$ terms vanish. Within the integrand this leads to the Eq. (9) vector differential equation for the Lagrange multiplier variables, which are sometimes alternatively called influence, auxiliary, complementary or co-state variables.

$$\underline{\dot\lambda}^T = -\frac{\partial H}{\partial\underline{x}} = -\frac{\partial L}{\partial\underline{x}} - \underline{\lambda}^T\frac{\partial\underline{f}}{\partial\underline{x}} \qquad\qquad (9)$$

If the initial system state is fully specified as in Eq. (3) making $\underline{\delta x}(t_0)=\underline{0}$, then the second term of Eq. (8) vanishes regardless of $\underline{\lambda}(t_0)$. If the final state is not specified thus making

$\underline{\delta x}(t_f) \neq \underline{0}$, then in order for the first term of Eq. (8) to vanish its coefficient must equal zero implying the Eq. (10) final condition for Eq. (9):

$$\underline{\lambda}(t_f) = \left(\frac{\partial \varphi}{\partial \underline{x}} \right)^T \Bigg|_{t_f} .$$

(10)

Alternatively, if the final state is specified as in Eq. (11),

$$\underline{x}(t_f) = \underline{x}_f ,$$

(11)

so that $\underline{\delta x}(t_f) = \underline{0}$, then there is no specified boundary condition for Eq. (9), but Eq. (1) has all state variables specified at both initial and final times. Either way, the first three terms of Eq. (8) are zero, leaving only the fourth term:

$$\delta J = \int_{t_0}^{t_f} \left[\frac{\partial H}{\partial \underline{u}} \delta \underline{u} \right] dt .$$

(12)

In cases where the control is unconstrained, the control variations $\underline{\delta u}(t)$ can be either positive or negative so if $\delta J \neq 0$, then it becomes impossible to ensure either Eq. (7a) or (7b). This implies when considering small amplitude variations for either maximizing or minimizing unconstrained control situations, $\delta J = 0$. Thus, the optimal solution with unconstrained control is found by setting $\frac{\partial H}{\partial \underline{u}} = 0$, solving the resulting algebraic equations for $\underline{u}(t)$ in terms of $\underline{x}(t)$ and $\underline{\lambda}(t)$, and then numerically solving the system of these relations in conjunction with Eqs. (2), (3), (9) and the final boundary condition, which is either Eq. (10) if the final state is unspecified or Eq. (11) if the final state is specified.

If instead, the control is constrained by Eq. (1), then admissible control variations must be restricted so that $\delta \underline{u}(t) + \underline{u}(t) \in U$, which complicates analysis by making $\delta \underline{u}$ depend on $\underline{u}$

through a set of inequalities. These can be worked through, and a solution can be found for a control optimal with respect to small variations.

However, the Pontryagin's Maximum Principle (PMP) theorem may instead be applied for either constrained or unconstrained situations, choosing $\underline{u}$ consistent with Eq. (1) to pointwise maximize (or minimize) H to find the optimal solution. Where it applies, the PMP is more powerful since it optimizes over both small and large control variations. Either way, determining the optimal control requires simultaneous solution of the two vector differential equations, Eqs. (2) and (9), coupled through algebraic equations derived from $\dfrac{\partial H}{\partial \underline{u}} = 0$ or from the PMP and with split boundary conditions given by Eq. (3) at the initial time and by either Eq. (10) or Eq. (11) at the final time.

Russian mathematician Pontryagin together with his students derived the maximum (or minimum) principle in 1956, expanding it into a 1961 book subsequently translated into English.[6,7] The PMP supposes that an admissible control history $\underline{u}_{opt}(t)$ obeying Eq. (1), along with its associated state history $\underline{x}_{opt}(t)$ satisfying Eq. (2) with boundary condition Eq. (3), is known to be optimal in maximizing (or minimizing) Eq. (4), i.e., that Eq. (7) is known to hold for all admissible variation pairs. The PMP states that then a nonzero vector function, $\underline{\lambda}_{opt}(t)$, satisfying Eq. (9), exists with the property that separately at each time point, the optimal control $\underline{u}_{opt}(t)$ necessarily maximizes (or minimizes) pointwise the Hamiltonian H function with respect to all other possible values of the control at that same time point. Stated mathematically, then $\underline{\lambda}_{opt}(t)$, satisfying Eq. (9), exists with the property that:

$$\underline{u}_{opt}(t) = \underline{\upsilon}^*(t) \in U \qquad \forall t \in [t_0, t_f] \qquad \text{where}$$

for a maximizing optimal control:

$$H\left(\underline{x}_{opt}(t), \underline{\upsilon}^*(t), \underline{\lambda}_{opt}(t), t\right) = \max_{\underline{\upsilon} \in U}\left(H\left(\underline{x}_{opt}(t), \underline{\upsilon}, \underline{\lambda}_{opt}(t), t\right)\right) \quad \text{(a)}$$

for a minimizing optimal control:

$$H\left(\underline{x}_{opt}(t), \underline{\upsilon}^*(t), \underline{\lambda}_{opt}(t), t\right) = \min_{\underline{\upsilon} \in U}\left(H\left(\underline{x}(t)_{opt}, \underline{\upsilon}, \underline{\lambda}_{opt}(t), t\right)\right) \quad \text{(b)} \tag{13}$$

Pontryagin's proof for general finite-dimensional nonlinear dynamic systems was only for necessary conditions. If a control optimizing Eq. (4) exists, then it necessarily must satisfy Eqs. (1), (2), (3), (9) and (13). Pontryagin did not prove the converse that merely satisfying Eqs. (1), (2), (3), (9) and (13) with appropriate final boundary condition Eq. (10) or Eq. (11) would be sufficient to conclude that a control $\underline{u}(t)$ optimizes Eq. (4).

On the other hand, from a practical engineering perspective, any control solution satisfying Eqs. (1), (2), (3), (9), and (13) with appropriate final boundary condition Eq. (10) or Eq. (11) may be of interest even in cases without PMP sufficiency. Such a solution satisfies all mission requirements, and small admissible variations cannot improve on its performance. Concerns about PMP sufficiency arise because in some degenerate situations there may not be any optimal control solution (e.g., there is no smallest positive number) and in other situations there may be multiple control histories that are locally optimal. Pontryagin's book discusses sufficiency in its Chap. 1, Sec. 4, advancing an argument that optimal solutions are topologically isolated (which almost proves sufficiency) and that if it can be demonstrated by other means that an optimal control exists, then the optimal control must be among the solutions of Eqs. (1), (2), (3), (9), (13) and either (10) or (11). If also the solutions to this system can be proven to be unique, then PMP sufficiency follows. Pontryagin's Chap. 4 discusses minimum-time optimal control of linear systems, a special class for which existence and uniqueness theorem proofs are presented.

After Pontryagin's book appeared, further progress was made by others on sufficiency as summarized in a 1977 paper,[8] which also further extended sufficiency theorem proofs to situations including state variable constraints. It recounted that a 1966 paper[9] had proven that PMP sufficiency follows for any Hamiltonian function that is concave in both control and state variables for a maximization case or convex for minimization. Since concavity and convexity each include linearity, the PMP is necessary and sufficient in a wide variety of cases. A more general result in 1968 established that concavity (or convexity) with respect to the control variables is not essential, showing that if the Hamiltonian function is first optimized over the admissible control set U, then PMP sufficiency follows if the resulting optimized function is concave (convex) with respect to state variables only.[10]   Much of optimal control theory for finite dimensional systems has been incorporated in accessible textbooks.[11]

II. OPTIMAL SHIELD DESIGN

Behaviors of shields are governed by the physics of radiation transport. For neutral particles such as neutrons or gamma rays, these are simplified by the fact that particle paths between interactions with atomic nuclei are straight lines unaffected by electromagnetic fields. Conservation of particles implies that the net particle addition rate matches the rate of change in local particle density N, but the convective derivative $\dfrac{DN}{Dt} \equiv \dfrac{\partial N}{\partial t} + \vec{v} \bullet \nabla N$ where $\vec{v}$ is particle velocity must be used to express conservation in a stationary frame. It is customary to express particle velocity as the product of speed and a direction-of-motion unit vector, $\vec{v} \equiv v\hat{\Omega}$, to replace particle density by angular flux, the product of particle speed and density, $\psi \equiv vN$, and to use particle energy E as an independent variable instead of particle speed which then

depends on energy and particle rest mass m as $v = \dfrac{c\sqrt{E^2 + 2mc^2E}}{E + mc^2}$, where c is the speed of light. With these customary variables, conservation is expressed by the time-dependent BTE as follows:

$$\frac{1}{v}\frac{\partial}{\partial t}\psi(\vec{r},\hat{\Omega},E,t) + \hat{\Omega}\bullet\nabla\psi(\vec{r},\hat{\Omega},E,t) + \sigma_t(\vec{r},E)\psi(\vec{r},\hat{\Omega},E,t) = q_t(\vec{r},\hat{\Omega},E,t)$$

$$\begin{pmatrix}\text{Rate of}\\ \text{Increase}\end{pmatrix} + \begin{pmatrix}\text{Rate of Out -}\\ \text{Convection}\end{pmatrix} + \begin{pmatrix}\text{Total Removal Rate}\\ \text{by Reactions}\end{pmatrix} = \begin{pmatrix}\text{Total Source}\\ \text{Addition Rate}\end{pmatrix} \qquad\text{.(14)}$$

In Eq. (14), the angular flux distribution's single scalar variable, ψ, is modeled as being dependent on seven independent scalar variables represented by position $\vec{r}$, particle direction $\hat{\Omega}$, particle energy E, and time, t. The macroscopic total cross section σ_t for all reactions removing particles from a particular energy and direction depends on incident particle energy and also varies with position if shield material composition is nonuniform.

It should be noted that the $\hat{\Omega}\bullet\nabla\psi$ term in Eq. (14) is simply the directional derivative of ψ in the $\hat{\Omega}$ direction and that no other terms explicitly refer to other directions, even via differential or integral operators. Therefore, if its right hand side were fully known then one could express the exact angular flux solution to Eq. (14) in closed analytical form by integrating along lines, a fact forming the basis for integral transport methods[12] and also related to Monte Carlo solution methods. However, in most situations the right hand side's total source also depends on the angular flux function to be determined. The total particle source addition rate q_t can be decomposed into the sum of source contributions from any external fixed source specified, the inherent inscattering source, and any other flux-dependent source types present that for gamma rays would include photons released by radiative neutron captures or for neutrons would include fission sources. Thus, for neutrons,

$$q_T(\vec{r},\hat{\Omega},E,t) \;=\; q_{ex}(\vec{r},\hat{\Omega},E,t) \;+\; q_s(\vec{r},\hat{\Omega},E,t) \;+\; q_f(\vec{r},E,t)$$

$$\begin{pmatrix} \text{Total} \\ \text{Source} \\ \text{Addition} \\ \text{Rate} \end{pmatrix} = \begin{pmatrix} \text{External} \\ \text{Source} \\ \text{Addition} \\ \text{Rate} \end{pmatrix} + \begin{pmatrix} \text{InScattering} \\ \text{Addition} \\ \text{Rate} \end{pmatrix} + \begin{pmatrix} \text{Fission} \\ \text{Daughter} \\ \text{Neutron} \\ \text{Addition} \\ \text{Rate} \end{pmatrix}. \tag{15}$$

The scattering source term introduces complexity in that it involves integration over energy of the incident particles being scattered, and over the deflection angle between incoming and outgoing directions in a formula which requires use of double-differential cross sections:

$$q_s(\vec{r},\hat{\Omega},E,t) = \int dE' \oiint d\Omega' \sigma_s(E' \to E, \hat{\Omega}' \bullet \hat{\Omega} \equiv \mu) \psi(\vec{r},\hat{\Omega}',E',t). \tag{16}$$

Although radiation transport physics is fundamentally described as the time-dependent solution of Eq. (14), the transient angular flux response to a suddenly applied radiation source step function of time typically varies on a nano-second time scale and reaches a steady-state condition in much less than a second. Details of such fast time responses are not of interest to shield designers concerned with steady behavior. After the brief initial transient, the time derivative term in Eq. (14) becomes insignificant so it is omitted altogether in the time-independent version of the BTE normally used to represent shielding physics:

$$\hat{\Omega} \bullet \nabla \psi(\vec{r},\hat{\Omega},E) + \sigma_t(\vec{r},E) \psi(\vec{r},\hat{\Omega},E) =$$
$$\int dE' \oiint d\Omega' \sigma_s(E' \to E, \hat{\Omega}' \bullet \hat{\Omega} \equiv \mu) \psi(\vec{r},\hat{\Omega}',E') + q(\vec{r},\hat{\Omega},E), \tag{17}$$

where the retained $q(\vec{r},\hat{\Omega},E)$ term represents any particle sources present in addition to inscattering.

Shielding optimization herein applies an approach similar to optimal control theory but different in that the independent scalar variable t (time) of optimal control theory is replaced by six independent scalar variables: $(\vec{r},\hat{\Omega},E)$ representing position, particle direction and particle

energy; the control vector $\underline{u}(t) \in \mathfrak{R}^m$ is replaced by the shield design $\underline{u}(\vec{r}) \in \mathfrak{R}^m$; and the state

vector $\underline{x}(t) \in \mathfrak{R}^n$ is replaced by the angular flux function $\psi(\vec{r}, \hat{\Omega}, E) \in \mathfrak{R}^+$. System dynamics is

replaced by Eq. (17), the time-independent BTE, which is a linear first order PDE also including

linear integral terms for scattering.

Shielding design problem specifications include a radiation source description, a definition of

the "detector" regions in which radiation dose rates must be limited along with their limiting

dose rate values, and a definition of which shielding materials are allowed and where.

II.A. Shielding Definitions and Notation

As shown in Fig. 1, the convex spatial domain $\mathcal{D} \subset \mathfrak{R}^3$ surrounded by vacuum contains a

radiation source and subregions $\mathcal{D}_\alpha \subset \mathcal{D}$ indexed by $\alpha \in A$ in which the dose rate must be

limited. The scalar function $q(\vec{r}, \hat{\Omega}, E)$ specifies the radiation source distribution per unit solid

angle and energy at location $\vec{r}$ and particle energy E.

Table I is an example of the specified menu of allowed shielding materials which may be

used in a particular shielding design problem. With this menu represented as a sequenced list of

m allowable materials, the volume fraction in the shield of material number i, $i \in I$, where I is

the m-component material index set, is the scalar function of position, $u_i(\vec{r})$. A complete design

of a shield is thus given by specifying all m of these volume fraction scalar functions of position,

or equivalently, by the single m-vector function of position:

$$\underline{u}(\vec{r}) \equiv \left(u_1(\vec{r}) \quad u_2(\vec{r}) \quad \dots \quad u_m(\vec{r}) \right)^T .$$

(18)

$\mathscr{R}^3$

Radiation Source q

Shielding

Solve for Angular Flux ψ

$\mathscr{D}_1$

$\mathscr{D}_\alpha$
Doserate Constrained Sub-Regions

$\mathscr{D}_2$

$\mathscr{D}$

Convex Domain

Boundary
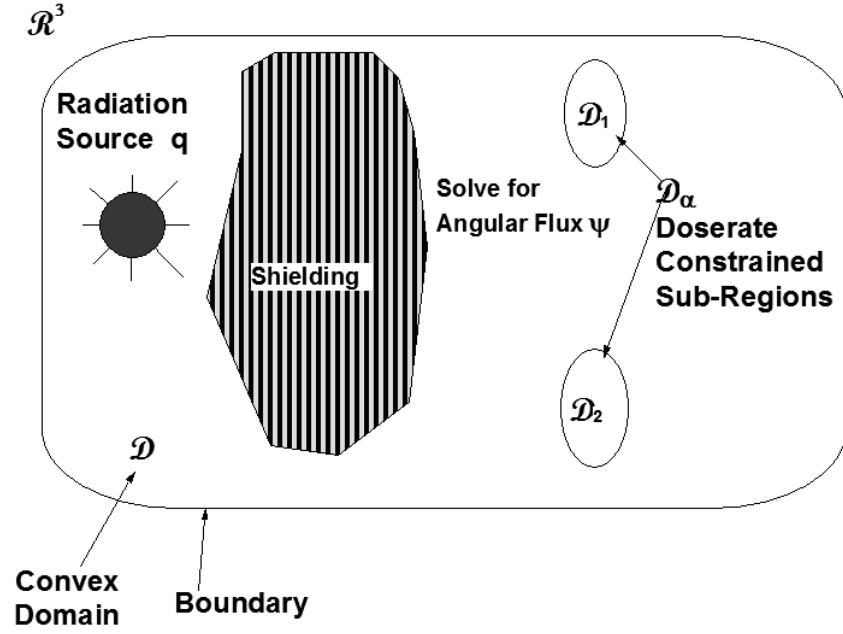
Fig.1: Shielding Problem Diagram

TABLE I

An Example of a Materials Menu

Materials Menu

#	material	cost/ volume	macroscopic total cross section	macroscopic scattering cross section
1	tungsten	c_1	$\sigma_t^{(1)}(E)$	$\sigma_s^{(1)}(\hat{\Omega}'\to\hat{\Omega}, E'\to E)$
2	boron-10	c_2	$\sigma_t^{(2)}(E)$	$\sigma_s^{(2)}(\hat{\Omega}'\to\hat{\Omega}, E'\to E)$
3	beryllium	c_3	$\sigma_t^{(3)}(E)$	$\sigma_s^{(3)}(\hat{\Omega}'\to\hat{\Omega}, E'\to E)$
4	iron	c_4	$\sigma_t^{(4)}(E)$	$\sigma_s^{(4)}(\hat{\Omega}'\to\hat{\Omega}, E'\to E)$
5	polyethylene	c_5	$\sigma_t^{(5)}(E)$	$\sigma_s^{(5)}(\hat{\Omega}'\to\hat{\Omega}, E'\to E)$
6	lead	c_6	$\sigma_t^{(6)}(E)$	$\sigma_s^{(6)}(\hat{\Omega}'\to\hat{\Omega}, E'\to E)$
7	lithium-6 hydride	c_7	$\sigma_t^{(7)}(E)$	$\sigma_s^{(7)}(\hat{\Omega}'\to\hat{\Omega}, E'\to E)$
8	uranium-238	c_8	$\sigma_t^{(8)}(E)$	$\sigma_s^{(8)}(\hat{\Omega}'\to\hat{\Omega}, E'\to E)$
9	graphite	c_9	$\sigma_t^{(9)}(E)$	$\sigma_s^{(9)}(\hat{\Omega}'\to\hat{\Omega}, E'\to E)$
10	lithium-7 hydride	c_{10}	$\sigma_t^{(10)}(E)$	$\sigma_s^{(10)}(\hat{\Omega}'\to\hat{\Omega}, E'\to E)$
11	boron-10 carbide	c_{11}	$\sigma_t^{(11)}(E)$	$\sigma_s^{(11)}(\hat{\Omega}'\to\hat{\Omega}, E'\to E)$
12	water	c_{12}	$\sigma_t^{(12)}(E)$	$\sigma_s^{(12)}(\hat{\Omega}'\to\hat{\Omega}, E'\to E)$
...	...	...	...	...
m	material # m	c_m	$\sigma_t^{(m)}(E)$	$\sigma_s^{(m)}(\hat{\Omega}'\to\hat{\Omega}, E'\to E)$

Volume fractions are by definition constrained as $0 \le u_i(\vec{r}) \le 1$, $i \in I$ and also by $\sum_{i \in I} u_i(\vec{r}) \le 1$. This is stated succinctly as $\underline{u}(\vec{r}) \in U$, where $U \subset \mathfrak{R}^m$ is the set equivalent to those definitional constraints. For example, this set can be visualized for m=3 by Fig. 2.
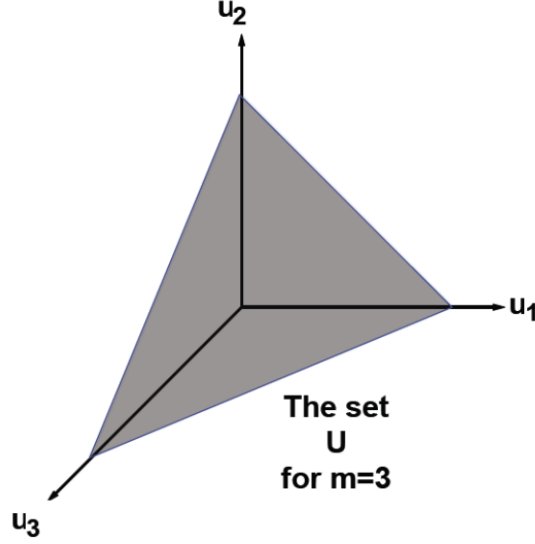
Fig. 2: Control Constraint Set U for m=3

In practical shielding design situations, there may be other position-dependent constraints further restricting the materials choices in portions of the spatial domain, here stated as follows:

$$\underline{u}(\vec{r}) \in U_c(\vec{r}) \subset U ,$$

(19)

where, $U_c(\vec{r})$ is the specified position-dependent set of material choice options under the control of the shielding designer.

Total materials cost C is modeled in terms of volumetric cost rates c_i assigned for each material $i \in I$. Defining the materials cost vector as $\underline{c} = (c_1 \quad c_2 \quad ... \quad c_m)^T$, the total materials cost is

$$C = \underline{c}^T \iiint_{\mathcal{D}} \underline{u}(\vec{r}) dV$$

.

(20)

Scattering and total cross sections are also vectorized as

$$\underline{\sigma}_t(E) \equiv \left(\sigma_t^{(1)}(E) \quad \sigma_t^{(2)}(E) \quad \cdots \quad \sigma_t^{(m)}(E) \right)^T .$$

(21)

and

$$\underline{\sigma}_s\left(\hat{\Omega}' \to \hat{\Omega}, E' \to E\right) \equiv \begin{pmatrix} \sigma_s^{(1)}\left(\hat{\Omega}' \to \hat{\Omega}, E' \to E\right) \\ \sigma_s^{(2)}\left(\hat{\Omega}' \to \hat{\Omega}, E' \to E\right) \\ \vdots \\ \sigma_s^{(m)}\left(\hat{\Omega}' \to \hat{\Omega}, E' \to E\right) \end{pmatrix} .$$

(22)

The BTE for angular flux, $\psi\left(\vec{r}, \hat{\Omega}, E\right)$, is then written in vectorized form as:

$$\hat{\Omega} \bullet \nabla \psi\left(\vec{r}, \hat{\Omega}, E\right) + \underline{\sigma}_t^T(E)\underline{u}(\vec{r})\psi\left(\vec{r}, \hat{\Omega}, E\right) =$$

$$\oiint d\Omega' \int_0^\infty dE' \underline{\sigma}_s^T\left(\hat{\Omega}' \to \hat{\Omega}, E' \to E\right)\underline{u}(\vec{r})\psi\left(\vec{r}, \hat{\Omega}', E'\right) + q\left(\vec{r}, \hat{\Omega}, E\right) \quad \text{(a)}$$

$$\text{for} \quad \vec{r} \in \mathcal{D}, \underline{u}(\vec{r}) \in U_c(\vec{r}) \subset U \quad \text{(b)}$$

,

(23)

with boundary condition

$$\psi\left(\vec{r}, \hat{\Omega}, E\right)\Big|_{\substack{\vec{r} \in \partial\mathcal{D} \\ \hat{n}\bullet\hat{\Omega}<0}} = 0$$

,

(24)

where $\hat{n}$ is the outer-directed unit normal on the domain boundary $\partial\mathcal{D}$. Eqs. (21) through (24) can describe neutron or gamma ray transport which also may be combined by redefining E as a two-element vector. The dose rates D_α, within regions $\mathcal{D}_\alpha$, are the response of an omnidirectional detector separable in energy and position. Its energy weighting function $w(E)$ takes into account tissue absorption and quality factors. Its position weighting function, $W_\alpha(\vec{r})$, is normalized as

$$\iiint_{\mathcal{D}_\alpha} W_\alpha(\vec{r})dV = 1$$

.

(25)

Then the dose rates are:

$$D_\alpha = \iiint\limits_{\mathcal{D}_\alpha} dV \oiint d\Omega \int\limits_0^\infty dE\, W_\alpha(\vec{r}) w(E) \psi(\vec{r},\hat{\Omega},E)$$

(26)

Radiation dose rate constraints required of the shielding design are:

$$D_\alpha \le D_\alpha^{MAX} \quad \forall\, \alpha \in A ,$$

(27)

where D_α^{MAX} values are specified requirements.

II.B. Optimal Shield Design Problem Statement

The shield cost optimization design problem is to choose $\underline{u}(\vec{r}) \in U_c(\vec{r})$ in accordance with the Eq. (19) constraint so that the functional C in Eq. (20) is minimized while Eqs. (23), (24), (26) and (27) are satisfied.

II.C. Derivation of Shield Optimality Conditions

Eq. (20) is first augmented by adding two zero value terms:

$$C = \underline{c}^T \iiint\limits_{\mathcal{D}} dV\, \underline{u}(\vec{r}) + \sum_{\alpha \in A} \tau_\alpha \left(D_\alpha - D_\alpha^{MAX} \right)$$
$$+ \iiint\limits_{\mathcal{D}} dV \oiint d\Omega \int\limits_0^\infty dE\, \underline{\lambda}(\vec{r},\hat{\Omega},E) \left[\begin{pmatrix} -\underline{\sigma}_t^T(E)\psi(\vec{r},\hat{\Omega},E) + \\ \left(\oiint d\Omega' \int\limits_0^\infty dE'\, \underline{\sigma}_s^T(\hat{\Omega}' \to \hat{\Omega}, E' \to E)\psi(\vec{r},\hat{\Omega}',E') \right) \underline{u}(\vec{r}) \\ + q(\vec{r},\hat{\Omega},E) - \hat{\Omega} \bullet \nabla \psi(\vec{r},\hat{\Omega},E) \end{pmatrix} \right].$$

(28)

The first added term includes Kuhn-Tucker (KT) multipliers τ_α having the following definition:

$$\tau_\alpha \begin{cases} = 0 \ \ \text{if} \ \ D_\alpha - D_\alpha^{MAX} < 0 \\ \ge 0 \ \ \text{if} \ \ D_\alpha - D_\alpha^{MAX} = 0 \end{cases} .$$

(29)

In nonlinear optimization theory, the KT theorem[13] (1951) defines corner conditions for inequality constraints. It considers maximization of a continuously differentiable scalar objective function of n variables $h(\underline{x})$ [or of $-h(\underline{x})$ for minimization] within a region defined as the set intersection of m inequalities $\{g_\alpha(\underline{x}) \le 0\}_{\alpha=1}^{\alpha=m}$, where the g_α constraint functions are also continuously differentiable. The theorem asserts that $\left(\nabla h(\underline{x}) - \sum_{\alpha=1}^{\alpha=m} \tau_\alpha(\underline{x}) g_\alpha(\underline{x})\right)\Big|_{\underline{x}=\underline{x}^*} = \underline{0}$ holds at any local maximum point $\underline{x}^*$, where the undetermined nonnegative KT multiplier functions $\tau_\alpha(\underline{x})$ are defined such that they form identically zero products when multiplied by their associated constraint functions, i.e., $\tau_\alpha(\underline{x}) \ge 0$ and $\tau_\alpha(\underline{x}) g_\alpha(\underline{x}) = 0$. Interpreting the KT theorem, for interior points where $g_\alpha(\underline{x}) > 0 \quad \forall \alpha$, the KT theorem returns the familiar result that all first derivatives are zero at a maximum. On single active constraint portions of the region's boundary [where only one $g_\alpha(\underline{x})$ constraint function is zero], the KT theorem states that the objective function's gradient direction at a maximum point exactly matches the active constraint function's gradient direction there. At corner points of the region boundary where two or more constraints are active, the KT theorem states that the gradient direction of the objective function lies within the set of directions that can be formed as weighted sums of the gradients of the active constraint functions using nonnegative weighting coefficients. The KT theorem provides only necessary conditions for an optimum, but they become sufficient conditions also if the objective function is concave and the constraint functions are all convex.

Use here of KT multipliers is to multiply functions which are not known in advance but instead result from solving PDEs. Otherwise their use is conventional. It should be noted that Eqs. (27) and (29) together guarantee that the first added term in Eq. (28) is identically zero.

The second added term, including the Lagrange multiplier function $\lambda \equiv \lambda\left(\vec{r},\hat{\Omega},E\right)$, is identically

zero because of Eq. (23). It is useful to define a Hamiltonian function via Eq. (30):

$$H = \underline{c}^T \underline{u}(\vec{r}) + \oiint d\Omega \int_0^\infty dE \left(\begin{array}{l} \lambda\left(\vec{r},\hat{\Omega},E\right) \times \\ \left[\left[\left(-\underline{\sigma}_t^{\ T}(E)\psi\left(\vec{r},\hat{\Omega},E\right) + \right. \right. \\ \left. \left. \oiint d\Omega' \int_0^\infty dE' \underline{\sigma}_s^{\ T}\left(\hat{\Omega}' \to \hat{\Omega}, E' \to E\right)\psi\left(\vec{r},\hat{\Omega}',E'\right) \right)\underline{u}(\vec{r}) \right] \\ \left. + q\left(\vec{r},\hat{\Omega},E\right) \right] \\ + \sum_{\alpha \in A} \tau_\alpha W_\alpha(\vec{r})w(E)\psi\left(\vec{r},\hat{\Omega},E\right) \end{array} \right) . \tag{30}$$

The next step swaps the dummy integration variables for scattering and chooses the arbitrary

Lagrange multipliers to simplify Eq. (30), as follows:

$$\lambda\left(\vec{r},\hat{\Omega},E\right) \equiv \sum_{\alpha \in A} \tau_\alpha \psi_\alpha^{\ *}\left(\vec{r},\hat{\Omega},E\right), \tag{31}$$

where the $\psi_\alpha^{\ *}$ functions are chosen to satisfy normalized adjoint flux equations for each detector

region, dependent only on the design, $\underline{u}(\vec{r})$, i.e.,

$$-\hat{\Omega} \bullet \nabla \psi_\alpha^{\ *}\left(\vec{r},\hat{\Omega},E\right) + \underline{\sigma}_t^{\ T}\underline{u}(\vec{r})\psi_\alpha^{\ *}\left(\vec{r},\hat{\Omega},E\right) =$$
$$\oiint d\Omega' \int_0^\infty dE' \underline{\sigma}_s^{\ T}\left(\hat{\Omega} \to \hat{\Omega}', E \to E'\right)\psi_\alpha^{\ *}\left(\vec{r},\hat{\Omega}',E'\right)\underline{u}(\vec{r}) + W_\alpha(\vec{r})w(E), \tag{32}$$

with boundary conditions:

$$\psi_\alpha^{\ *}\left(\vec{r},\hat{\Omega},E\right) = 0, \quad \vec{r} \in \partial\mathcal{D}, \hat{n} \bullet \hat{\Omega} > 0. \tag{33}$$

The Eq. (30) integral is then integrated by parts in three dimensions using Greens Theorem.

The surface terms vanish because of Eq. (24) and the Eq. (33) choice. Because of Eqs. (31) and

(32), Eq. (30) is simplified to the following:

$$H = \underline{c}^T \underline{u}(\vec{r}) + \sum_{\alpha \in A} \tau_\alpha \oiint d\Omega \int_0^\infty dE \left(\begin{array}{l} \psi_\alpha^*(\vec{r},\hat{\Omega},E) q(\vec{r},\hat{\Omega},E) + \psi(\vec{r},\hat{\Omega},E) \times \\ \left[\left[\begin{array}{l} \left(-\underline{\sigma}_t^T(E) \psi_\alpha^*(\vec{r},\hat{\Omega},E) + \\ \oiint d\Omega' \int_0^\infty dE' \underline{\sigma}_s^T(\hat{\Omega} \to \hat{\Omega}', E \to E') \psi_\alpha^*(\vec{r},\hat{\Omega}',E') \end{array} \right) \underline{u}(\vec{r}) \right] \\ + W_\alpha(\vec{r}) w(E) \end{array} \right] \end{array} \right) .$$ (34)

Eq. (34) could be written more compactly by defining a discriminant vector function of position which also depends on the entire flux and adjoint flux functions, as follow:

$$\underline{d}_\alpha(\vec{r}; \psi_\alpha^*, \psi) \equiv \oiint d\Omega \int_0^\infty dE \, \psi(\vec{r},\hat{\Omega},E) \left(\begin{array}{l} -\underline{\sigma}_t(E) \psi_\alpha^*(\vec{r},\hat{\Omega},E) + \\ \oiint d\Omega' \int_0^\infty dE' \underline{\sigma}_s(\hat{\Omega} \to \hat{\Omega}', E \to E') \psi_\alpha^*(\vec{r},\hat{\Omega}',E') \end{array} \right) .$$ (35)

The PMP is next invoked as a minimum principle. For the present shielding design case the PMP states that the condition for optimality of a particular design, $\underline{u}(\vec{r})$, is that its choice of $\underline{u}$ at each point $\vec{r} \in \mathcal{D}$ optimizes the Hamiltonian H at that point with respect to all other admissible choices of $\underline{u} \in U_c(\vec{r})$ at that point, using the optimal design's own angular flux and normalized adjoint angular flux functions to evaluate H. Thus, the optimality condition is as follows:

$$\forall \vec{r} \in \mathcal{D}, \qquad \underline{u}_{opt}(\vec{r}) = \underline{\upsilon}^*(\vec{r}) \text{ where } \underline{\upsilon}^*(\vec{r}) \text{ obeys}$$

$$\left(\left(\underline{c}^T + \sum_{\alpha \in A} \tau_\alpha \underline{d}_\alpha^T(\vec{r}; \psi_\alpha^*, \psi) \Big|_{opt} \right) \underline{\upsilon}^*(\vec{r}) \right) = \min_{\underline{\upsilon} \in U_c(\vec{r})} \left(\left(\underline{c}^T + \sum_{\alpha \in A} \tau_\alpha \underline{d}_\alpha^T(\vec{r}; \psi_\alpha^*, \psi) \Big|_{opt} \right) \underline{\upsilon} \right) .$$ (36)

In summary, the optimal shield design is the simultaneous solution to Eqs. (23), (24), (26), (27), (29), (32), (33), (35) and (36). For convenience, the optimal design equations are collected in Fig. 3.

(1) Forward Boltzmann Transport Equation System

$$\hat{\Omega} \bullet \nabla \psi\left(\vec{r},\hat{\Omega},E\right) + \underline{\sigma}_t^{~T}(E)\underline{u}(\vec{r})\psi\left(\vec{r},\hat{\Omega},E\right)$$

$$-\oiint d\Omega' \int_0^\infty dE' \underline{\sigma}_s^{~T}\left(\hat{\Omega}' \to \hat{\Omega}, E' \to E\right)\underline{u}(\vec{r})\psi\left(\vec{r},\hat{\Omega}',E'\right) - q\left(\vec{r},\hat{\Omega},E\right) = 0$$

$$\text{for}\ \ \vec{r} \in \boldsymbol{\mathcal{D}},\ \ \underline{u}(\vec{r}) \in U_c(\vec{r}) \subset U;\ \ \text{with}\ \psi\left(\vec{r},\hat{\Omega},E\right)\Big|_{\substack{\vec{r} \in \partial\boldsymbol{\mathcal{D}} \\ \hat{n}\bullet\hat{\Omega}<0}} = 0$$

(23 & 24)

(2) Dose rate Constraints and Evaluations, one for each dose rate constraint location:

$$D_\alpha = \iiint_{\boldsymbol{\mathcal{D}}_\alpha} dV \oiint d\Omega \int_0^\infty dE\, W_\alpha(\vec{r})w(E)\psi\left(\vec{r},\hat{\Omega},E\right)$$

(26)

$$D_\alpha \le D_\alpha^{MAX} \quad \forall\, \alpha \in A$$

(27)

(3) Kuhn-Tucker Multiplier Constraints, one for each dose rate constraint location:

$$\forall\, \alpha \in A \qquad \tau_\alpha \begin{cases} = 0\ \text{if}\ D_\alpha - D_\alpha^{MAX} < 0 \\ \ge 0\ \text{if}\ D_\alpha - D_\alpha^{MAX} = 0 \end{cases}$$

(29)

(4) Normalized Component Adjoint Boltzmann Transport Equation Systems, one for each dose rate constraint location:

$$-\hat{\Omega} \bullet \nabla \psi_\alpha^*\left(\vec{r},\hat{\Omega},E\right) + \underline{\sigma}_t^{~T}\underline{u}(\vec{r})\psi_\alpha^*\left(\vec{r},\hat{\Omega},E\right) =$$

$$\oiint d\Omega' \int_0^\infty dE' \underline{\sigma}_s^{~T}\left(\hat{\Omega} \to \hat{\Omega}', E \to E'\right)\psi_\alpha^*\left(\vec{r},\hat{\Omega}',E'\right)\underline{u}(\vec{r}) + W_\alpha(\vec{r})w(E)$$

for $\forall\, \alpha \in A$

$$\vec{r} \in \boldsymbol{\mathcal{D}}, \underline{u}(\vec{r}) \in U_c(\vec{r}) \subset U\ \ \text{and with}\ \psi_\alpha^*\left(\vec{r},\hat{\Omega},E\right)\Big|_{\substack{\vec{r} \in \partial\boldsymbol{\mathcal{D}} \\ \hat{n}\bullet\hat{\Omega}>0}} = 0$$

(32 & 33)

(5) Vector Discriminant Functions, one for each dose rate constraint location:

$$\underline{d}_\alpha\left(\vec{r};\psi_\alpha^*,\psi\right) \equiv$$

$$\oiint d\Omega \int_0^\infty dE\, \psi\left(\vec{r},\hat{\Omega},E\right) \begin{pmatrix} -\underline{\sigma}_t^{~T}(E)\psi_\alpha^*\left(\vec{r},\hat{\Omega},E\right)+ \\ \oiint d\Omega' \int_0^\infty dE' \underline{\sigma}_s^{~T}\left(\hat{\Omega} \to \hat{\Omega}', E \to E'\right)\psi_\alpha^*\left(\vec{r},\hat{\Omega}',E'\right) \end{pmatrix}$$

(35)

(6) Pontryagin's Optimality Condition:

$$\forall \vec{r} \in \boldsymbol{\mathcal{D}}, \qquad \underline{u}_{opt}(\vec{r}) = \underline{\upsilon}^*(\vec{r})\ \text{where}\ \underline{\upsilon}^*(\vec{r})\ \text{obeys}$$

$$\left(\left(\underline{c}^T + \sum_{\alpha \in A}\tau_\alpha \underline{d}_\alpha^T\left(\vec{r};\psi_\alpha^*,\psi\right)\Big|_{opt}\right)\underline{\upsilon}^*(\vec{r})\right) = \min_{\underline{\upsilon} \in U_c(\vec{r})}\left(\left(\underline{c}^T + \sum_{\alpha \in A}\tau_\alpha \underline{d}_\alpha^T\left(\vec{r};\psi_\alpha^*,\psi\right)\Big|_{opt}\right)\underline{\upsilon}\right)$$

(36)

Fig. 3: Summary of the Optimal Shield Design Equations

Pontryagin's Maximum Principle Applicability

Although generation of a proof is beyond the scope of this paper, it is appropriate to discuss why the PMP, originally proven for systems governed by any finite number of ODEs in the single independent time variable, t, should be relied on to optimize the time-independent BTE system with its six independent scalar variables $(\vec{r}, \hat{\Omega}, E)$. The main issue for PMP application to the BTE is the extension from finite to infinite dimensional systems.

First, it should be noted that proving PMP applicability may not be crucial since shield designs can be evaluated on their own merits.

An intuitive PMP applicability argument starts with the time-dependent BTE, Eq. (14) in seven independent variables. This can be approximated by representing time-varying angular flux values on a discrete grid meshing the six independent scalar variables of spatial position within the domain, particle energy and particle direction, as in discrete ordinates methods. Differential operations with respect to position, energy and direction are replaced by finite difference or finite element approximations, and integrals are replaced by weighted sums. Time derivatives of the $n<\infty$ approximating angular flux nodal values are then expressed in terms of angular flux nodal values and shielding cross sections on the spatial subgrid, thus replacing Eq. (14) for angular flux by a finite system of time-dependent ODEs for angular flux. It follows that the PMP applies to any such approximation. Well known finite difference[14] and finite element[15] proofs show that, for linear operators as in the BTE and even for nonlinear systems with smooth solutions, the approximation error of such finite grid models under certain restrictions converges to zero as the grids are refined to arbitrarily finer resolutions. Since the PMP holds for all such

approximations, it is plausible that it also holds for the BTE itself in the limit of zero approximation error.

After Pontryagin's book appeared, there was an explosion of mathematical interest in extending the PMP proof, which a recent paper[16] traces to 1961, citing early publications proving the PMP for special classes of infinite dimensional problems. However, PDE systems in optimal control problems are more diverse than ODE systems so even their taxonomic classification is significant. This area has by now generated a rich literature of its own, including many technical papers and even books filled with mathematical proofs for different classes of cases. It turns out that the PMP holds in many PDE cases, but there are other cases in which it does not hold. Theoretical investigations are continuing and the body of abstract literature continues to expand.

A 1968 seminal volume by Lions translated from French in 1971 (Ref. 17) focused on optimal control of second or higher order elliptic, parabolic or hyperbolic PDEs. The framing of these problems established terminology and methods of proof including, e.g., distinguishing between boundary control and what it called distributed control where the control is applied to the domain region's interior (as with optimal shielding design). A 1995 book[18] summarized much of the growing field's results. Results for optimal control of second-order PDEs have been extended[19,20] and so can be applied with confidence in engineering situations.

Another division in optimally controlled PDE problems is between evolution type problems in which one of several independent variables plays a distinctive time-like role, and Dieudonné-Rashevsky (DR) type problems in which the independent variables all have an equal rank and in which all derivatives are of first order.[21] The DR-type problems are of interest here because of their similarities to the time-independent BTE. In 1969 a proof of the PMP for DR-type problems in general was published.[22] However, it was later noticed that this proof had failed to

address some possible degenerate situations by simply assuming the existence of smooth solutions to a related Hamilton-Jacobi inequality in the neighborhood of an optimal control, which might be violated in certain DR-type systems. This gap was subsequently closed by reformulating the PMP as a so-called e-maximum principle.[23] The e-maximum principle was later shown to be automatically implied by augmenting the PMP with some regularity restrictions such as requiring the infinite dimensional state functions to obey a Lipschitz smoothness regularity constraint[24] or by requiring polyconvexity of the Hamiltonian.[25]

On the other hand, PMP proofs for DR-type systems do not exactly fit time-independent BTE systems whose first-order derivative terms are only with respect to the three spatial variables, while an integration operator acts with respect to the energy and direction variables. However, although the BTE's integrodifferential operator is not just a PDE, it is entirely linear, a fact which might provide a better basis for a BTE PMP proof. In 2010 Popescu[26] provided a general PMP proof for optimal control of time-evolution systems governed by general linear operators acting on Hilbert spaces representing both an infinite-dimensional state and an infinite-dimensional control. It included existence and uniqueness thus establishing sufficiency in addition to necessity. The proof's scope includes linear integrodifferential operators, but it may require some modification to precisely cover optimal design of shielding as governed by the time-independent BTE.

It should be mentioned that even if PMP sufficiency abstractly holds for optimal shielding design, computed designs may not be unique since practical numerical computations using floating point numbers will be insensitive to sufficiently small grain sizes in shielding mixtures. Thus, if uniqueness is essential to an algorithm, it may be necessary to impose a regularity restriction on $\underline{u}(\vec{r})$, although no such restrictions are used herein.

There are recent publications[27,28] applying optimal control methods with the BTE, albeit for an optimal boundary control application somewhat different from design of optimal radiation shielding.

In summary, producing a rigorous proof that the PMP holds for the BTE is beyond the scope of this paper, so the present optimal shielding design algorithm exploration proceeds based on the plausible belief that it applies.

II.D. Discussion of Optimality Conditions

Since the total cross sections include scattering, the components of discriminant vector functions $\underline{\boldsymbol{d}}_\alpha\left(\vec{r};\psi_\alpha^{\,*},\psi\right)$ as defined by Eq. (35) are nonpositive. Costs listed in the $\underline{c}$ vector are nonnegative as are the KT multipliers τ_α. Therefore, components of

$$\underline{b}(\vec{r}) \equiv \underline{c} + \sum_{\alpha \in A} \tau_\alpha \underline{\boldsymbol{d}}_\alpha\left(\vec{r};\psi_\alpha^{\,*},\psi\right)$$

(37)

may be negative or positive at different locations. At locations where all m components of $\underline{b}(\vec{r})$ are nonnegative, the optimal shield design minimizes $\underline{b}^T(\vec{r})\underline{u}(\vec{r})$ by setting $\underline{u}_{opt}(\vec{r}) = \underline{0}$ if $\underline{0} \in U_c(\vec{r})$ allows a void to be used there as the shielding material. In other locations where one or more components of $\underline{b}(\vec{r})$ are negative, the index of the most negative component specifies the optimal material choice. There can also be equally negative components of $\underline{b}(\vec{r})$ for which auction ties between different materials can be awarded arbitrarily within the Eq. (19) constraint.

Increasing the positive value of a KT multiplier τ_α cannot increase but may decrease the components of $\underline{b}(\vec{r})$ through Eq. (37), so the region in which some components of $\underline{b}(\vec{r})$ are negative can be expanded. This has the effect through Eq. (35) of adding shielding material at the expense of void regions, tending to reduce dose rates D_α. Thus, there is a continuous

relation between τ_α and D_α. The existence of this relation is fundamental to the optimal design algorithms discussed in the following. However, Eq. (29) does not allow positive τ_α values unless corresponding dose rates equal their maximum limits. Thus, optimal design solutions include precise τ_α values. Since Eq. (29) requires the corresponding τ_α to be zero for each dose rate strictly less than its maximum limit, it follows that Eq. (35) does not allow inactive constraints to influence the optimal design.

In theory, the optimal shield design equations provide a rigorous way to determine whether a particular radiation shield design is optimal. Using that design's $\underline{u}(\vec{r})$, solve for the angular flux and normalized component adjoint functions, calculate dose rates, then check whether Pontryagin's optimality condition is everywhere met for some set of KT multipliers conforming to dose rate-dependent positivity limitations. However, there is a practical impediment due to the extreme sharpness of the dose rate limitations' coupling with the KT multipliers. If the dose rate at any prescribed radiation constraint point is slightly less than the allowable dose rate there, even 10^{-100} Gy/yr less, then Eq. (29) requires the associated KT multiplier to be zero. The only permitted way for a particular multiplier to be positive is for the radiation at that location to precisely match the allowable dose rate there. On the other hand, if the radiation dose at that location is slightly greater than allowable, even 10^{-100} Gy/yr more, then that entire solution must be discarded as inadmissible since it violates the inequality in Eq. (27). Such sharp precision is not compatible with computer calculations using floating point numbers, so a modified approach is necessary.

III. OPTIMAL DESIGN ALGORITHM

Unfortunately Pontryagin's principle does not provide any algorithm to directly *find* the optimum design. A difficulty is that the logic is circular. To find the optimal shield design, $\underline{u}_{opt}(\vec{r})$ via Eq. (35), one must first compute $\underline{d}_\alpha\left(\vec{r}; \psi_\alpha^{\,*}, \psi\right)_{opt}$ via Eq. (36) for which one needs to know optimal angular flux ψ_{opt} and adjoint functions, $\left(\psi_\alpha^{\,*}\right)_{opt}$ that, in turn, depend on the optimal design, $\underline{u}_{opt}(\vec{r})$.

The heuristic iterative algorithm investigated here has two nested loops.   The sharpness of the coupling between dose rate limitations and KT multipliers is broken by passing fixed KT values to the inner loop which does not modify them. The inner loop ignores the Eq. (27) dose rate constraints but solves for an optimal shield design using an internal iteration scheme to break the otherwise circular logic concerning optimal flux functions. The inner loop returns to the outer loop with a shield design which is optimal for some set of allowable dose rates generally different from those actually specified.   The outer loop then modifies its KT parameters before again calling the inner loop in an algorithm eventually converging to the specified maximum allowed dose rates.   Thus, the inner loop iteratively modifies $\underline{u}(\vec{r})$ to converge to an optimal design for some set of D_α values, and the outer loop adjusts τ_α values to drive the D_α towards consistency with D_α^{MAX}.

The inner loop's iteration scheme is based on using old flux functions from its previous iteration's shield design. This would be a good approximation to Eq. (35) if angular flux functions changed only slightly between iterations but poor if changes are extensive. To help ensure this approximation is adequate, the inner loop only allows a small part of the shield to be

redesigned in a single iteration. Shield redesign modifications calculated from Eq. (35) using old flux functions are censored by limiting the redesign to a small, carefully chosen subset of the domain, $\mathcal{E}_k \subset \mathcal{D}$.

The inner loop is initialized with values for $\{\tau_\alpha\}_{\alpha \in A}$ and with an initial shield design $^{[k]}\underline{u}(\vec{r})$ for iteration count $k=0$. Eqs. (23), (32), and (33) are next solved for angular flux $^{[k]}\psi(\vec{r}, \hat{\Omega}, E)$ and normalized adjoint angular flux $^{[k]}\psi_\alpha^*(\vec{r}, \hat{\Omega}, E)$ functions, then Eqs. (36) and (37) are evaluated using these functions to determine $^{[k]}\underline{b}(\vec{r})$. Then instead of using Eq. (35), the algorithm evaluates Eq. (38) to determine the subsequent iteration's shield design:

$$\forall \vec{r} \in \mathcal{D}, \quad \underline{v}(\vec{r}) = \underline{\upsilon}^*(\vec{r}) \text{ where } \underline{\upsilon}^*(\vec{r}) \text{ obeys}$$
$$\left(^{[k]}\underline{b}^T(\vec{r}) \underline{\upsilon}^*(\vec{r}) \right) = \min_{\underline{\upsilon} \in U_c(\vec{r})} \left(^{[k]}\underline{b}^T(\vec{r}) \underline{\upsilon} \right)$$

then

$$^{[k+1]}\underline{u}(\vec{r}) = \begin{cases} \underline{v}(\vec{r}) & \forall \vec{r} \in \mathcal{E}_k \subset \mathcal{D} \\ ^{[k]}\underline{u}(\vec{r}) & \forall \vec{r} \in \mathcal{D} - \mathcal{E}_k \end{cases} \tag{38}$$

The Hamiltonian reduction improvement density between iterations [k] and [k+1] is then $^{[k]}\underline{b}^T(\vec{r}) \left(^{[k]}\underline{u}(\vec{r}) - \underline{v}(\vec{r}) \right)$ for locations within the redesign subset $\mathcal{E}_k \subset \mathcal{D}$. To maximize the iteration's improvement, an automated method sorts the $^{[k]}\underline{b}^T(\vec{r}) \left(^{[k]}\underline{u}(\vec{r}) - \underline{v}(\vec{r}) \right)$ values over the entire domain and then sets a threshold ε_k based on the highest percentile of the sorted volumetric improvements. Here, ε_k is a positive scalar threshold value for deciding whether departures from optimality are severe enough to merit changing material at a location during the current redesign iteration. The domain subset $\mathcal{E}_k \subset \mathcal{D}$ is then chosen based on its providing the largest projected Hamiltonian reductions, equivalent to the following redesign subset definition:

$$\mathcal{E}_k = \left\{ \vec{r} \in \mathcal{D} : \quad ^{[k]}\underline{b}^T(\vec{r}) \left(^{[k]}\underline{u}(\vec{r}) - \underline{v}(\vec{r}) \right) \geq \varepsilon_k \right\} \tag{39}$$

A flow chart of the inner loop appears as Fig. 4. It is simplified in that it does not show the sorting needed to determine $\mathcal{E}_k \subset \mathcal{D}$ nor does it show loop termination decisions based on recognition of convergence to an optimal design.

The outer loop appears as Fig. 5. Its purpose is to find, through iterative search calculations, the proper numerical values for the KT parameter set, $\{\tau_\alpha\}_{\alpha \in A}$ that result in dose rates, $\{D_\alpha\}_{\alpha \in A}$, consistent with the specified dose rate constraint set, $\{D_\alpha^{MAX}\}_{\alpha \in A}$. The outer loop's scheme is based on there being a continuous but unknown functional relationship, $\underline{f}$, between the ordered set of KT parameters passed to the inner loop, $\{\tau_\alpha\}_{\alpha \in A}$, and the resulting ordered set of dose rates returned from the inner loop, $\{D_\alpha\}_{\alpha \in A}$.

If the index set denoting dose rate constrained regions A in a shield design specification is finite with length n, the $\alpha \in A$ indices can be relabeled using integer numbers 1 through n then combined into n-dimensional vectors as in Eq. (40):

$$\underline{x} = (\tau_1, \tau_2, ..., \tau_n)^T \qquad \text{(a)}$$
$$\underline{y} = (D_1, D_2, ..., D_n)^T \qquad \text{(b)} \qquad\qquad (40)$$

Then the functional relationship is expressed by Eq. (41):

$$\underline{y} = \underline{f}(\underline{x}) \qquad\qquad (41)$$

In other shield design specifications where the index set is infinite, e.g., representing all points on a defined surface or in a defined volume, it may be true that in a particular solution only a finite number of points have dose rates equal to their maximum allowed values while the infinite number of other locations have dose rates less than their maximum allowed values. In such cases Eqs. (40) and (41) would still apply for the finite number of active variables.

Given Fixed $\{\tau_\alpha\}_{\alpha \in A}$

Given Initial Design ${}^{[0]}\underline{u}(\vec{r})$

Initialize loop counter: $k \leftarrow 0$

Increment loop count
$k \leftarrow k+1$

Find ${}^{[k]}\psi(\vec{r},\hat{\Omega},E)$ solving Forward Boltzmann Equation, i.e.:

$$\hat{\Omega} \bullet \nabla\left({}^{[k]}\psi(\vec{r},\hat{\Omega},E)\right) + \underline{\sigma}_t^{\,T}(E)\left({}^{[k]}\underline{u}(\vec{r})\right)\left({}^{[k]}\psi(\vec{r},\hat{\Omega},E)\right)$$

$$- \oiint d\Omega' \int_0^\infty dE' \underline{\sigma}_s^{\,T}\left(\hat{\Omega}' \to \hat{\Omega}, E' \to E\right)\left({}^{[k]}\underline{u}(\vec{r})\right)\left({}^{[k]}\psi(\vec{r},\hat{\Omega},E)\right) = q(\vec{r},\hat{\Omega},E)$$

over $\vec{r} \in \mathcal{D}$, with boundary condition : ${}^{[k]}\psi(\vec{r},\hat{\Omega},E)\Big|_{\substack{\vec{r} \in \partial\mathcal{D} \\ \hat{n} \bullet \hat{\Omega} < 0}} = 0$

Find ${}^{[k]}\psi_\alpha^{\,*}(\vec{r},\hat{\Omega},E)$ Solving Normalized Component Adjoint Boltzmann Equations, one for each dose rate constraint location

$$-\hat{\Omega} \bullet \nabla\left({}^{[k]}\psi_\alpha^{\,*}(\vec{r},\hat{\Omega},E)\right) + \underline{\sigma}_t^{\,T}\left({}^{[k]}\underline{u}(\vec{r})\right)\left({}^{[k]}\psi_\alpha^{\,*}(\vec{r},\hat{\Omega},E)\right)$$

$$- \oiint d\Omega' \int_0^\infty dE' \underline{\sigma}_s^{\,T}\left(\hat{\Omega} \to \hat{\Omega}', E \to E'\right)\left({}^{[k]}\psi_\alpha^{\,*}(\vec{r},\hat{\Omega}',E')\right)\left({}^{[k]}\underline{u}(\vec{r})\right) = W_\alpha(\vec{r})w(E)$$

$\forall \alpha \in A, \vec{r} \in \mathcal{D}$, with boundary conditions ${}^{[k]}\psi_\alpha^{\,*}(\vec{r},\hat{\Omega},E)\Big|_{\substack{\vec{r} \in \partial\mathcal{D} \\ \hat{n} \bullet \hat{\Omega} > 0}} = 0$

Calculate kth Set Discriminant Vector Functions of Position:

$${}^{[k]}\underline{d}_\alpha\left(\vec{r}; {}^{[k]}\psi_\alpha^{\,*}, {}^{[k]}\psi\right) = \oiint d\Omega \int_0^\infty dE \left({}^{[k]}\psi_\alpha^{\,*}(\vec{r},\hat{\Omega},E)\right) \times$$

$$\begin{pmatrix} -\underline{\sigma}_t^{\,T}(E)\left({}^{[k]}\psi(\vec{r},\hat{\Omega},E)\right) \\ + \oiint d\Omega' \int_0^\infty dE' \underline{\sigma}_s^{\,T}\left(\hat{\Omega}' \to \hat{\Omega}, E' \to E\right)\left({}^{[k]}\psi(\vec{r},\hat{\Omega}',E')\right) \end{pmatrix}$$

Calculate kth set of dose rates: ${}^{[k]}D_\alpha \equiv \int_0^\infty {}^{[k]}\phi\left(\vec{r}_\alpha^D, E\right)w(E)dE$

Calculate (k+1)st shield design

$\forall \vec{r} \in \mathcal{D}$, $\underline{v}(\vec{r}) = \underline{\upsilon}^*(\vec{r})$ where $\underline{\upsilon}^*(\vec{r})$ obeys $\left({}^{[k]}\underline{b}^T(\vec{r})\underline{\upsilon}^*(\vec{r})\right) = \min_{\underline{\upsilon} \in U_c(\vec{r})}\left({}^{[k]}\underline{b}^T(\vec{r})\underline{\upsilon}\right)$

and ${}^{[k]}\underline{b}(\vec{r}) \equiv \underline{c} + {}^{[k]}\underline{d}_\alpha\left(\vec{r}; {}^{[k]}\psi_\alpha^{\,*}, {}^{[k]}\psi\right)$ Then ${}^{[k+1]}\underline{u}(\vec{r}) = \begin{cases} \underline{v}(\vec{r}) & \forall \vec{r} \in \mathcal{E}_k \subset \mathcal{D} \\ {}^{[k]}\underline{u}(\vec{r}) & \forall \vec{r} \in \mathcal{D} - \mathcal{E}_k \end{cases}$

Fig. 4: Shield Redesign Inner Loop Basic Iteration Scheme

```
                    ┌─────────────┐
                    │    START    │
                    └──────┬──────┘
                           │        ┌──────────────────────────────────┐
                           │        │ N-vector of $\underline{y} = D_\alpha^{MAX}$ values │
                           │        └──────────────────────────────────┘
                           ▼
              ┌────────────────────────┐
              │   $\underline{f}(\underline{x}) = \underline{y}$      │            ┌──────────┐
──────────────▶│     NUMERICAL          │───────────▶│   EXIT   │
              │                        │            └──────────┘
              │   EQUATION SOLVER      │
              └───────────┬────────────┘
                          │        ┌─────────────────────────────────┐
                          │        │ $N$-vector of $\underline{x} = \tau_\alpha$ values │
                          │        └─────────────────────────────────┘
                          ▼
              ┌────────────────────────┐
              │   ENTIRE               │
              │                        │
              │   INNER LOOP           │
              │                        │
              │   (Figure 4)           │
              └───────────┬────────────┘
                          │        ┌─────────────────────────────────────────┐
                          │        │ $N$-vector of $\underline{f}(\underline{x}) = {}^{[\infty]}D_\alpha$ values │
                          └────────┴─────────────────────────────────────────┘
```
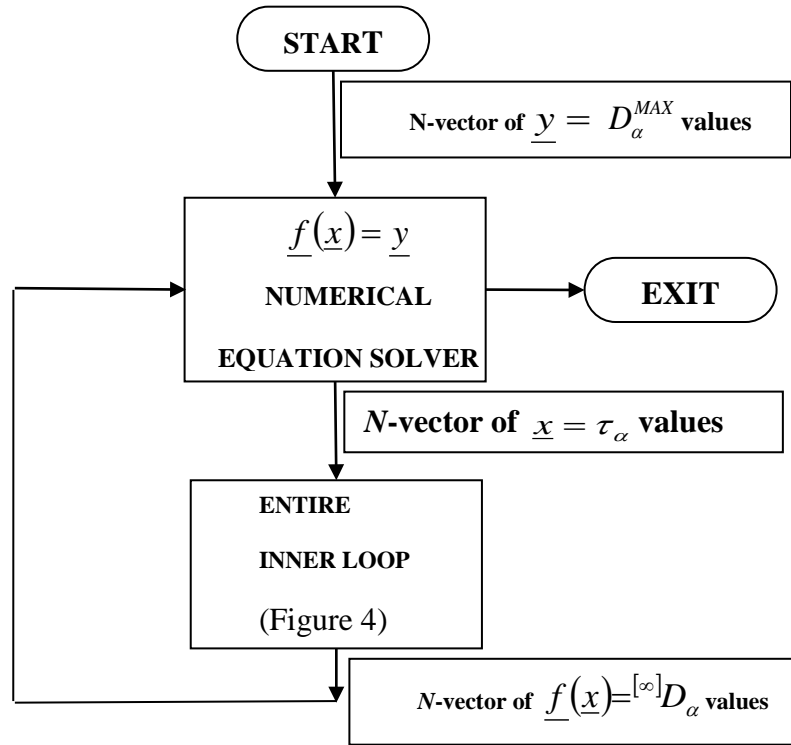
Fig. 5: Shield Redesign Outer Loop Basic Iteration Scheme

Although Eq. (41) determining dose rates, $\underline{y}$, from KT parameters, $\underline{x}$, can be evaluated by running the Fig. 4 inner loop algorithm to convergence, what is needed in the outer loop is the inverse calculation represented by Eq. (42).

$$\underline{x} = \underline{f}^{-1}(\underline{y})$$

(42)

However, there is no such algorithm to directly evaluate Eq. (42), so it is necessary to instead back into its solution using an iterative method. Various numerical algorithms have been developed to iteratively search for and find answers to such "numerical equation solving" problems and can be used here. For single-dimension searches appropriate for spherical shield designs, known standard numerical line search methods include bisection, Newton's method, various secant or chord methods, fixed-point Picard iteration schemes, and repetitive spline fits to previously acquired data that eventually learn the function's shape. Multidimensional searches can be performed using either sequences of one-dimensional line searches in successively different directions accomplished through any of these listed methods and connected together through Gauss-Jacobi or Gauss-Seidel iteration updates or, alternatively through multidimensional versions of either Newton's method, Picard iteration, Broyden's method, Powell's hybrid method or various homotopy continuation methods.

IV.   HEURISTIC ALGORITHM RESULTS

The optimal design algorithms were experimentally implemented for the spherically symmetric case in the SCALE (Standardized Computer Analyses for Licensing Evaluation) software system version 5.0 (Ref. 29) in a custom FORTRAN code module created by modifying SCALE's previously existing SAS1X control sequence to include them. As stated in the abstract of the SCALE Version 5 manual, "Since the initial release of SCALE in 1980, the code system has been widely used for evaluation of nuclear fuel facility and package designs." By now, the SCALE system has been extensively benchmarked and has become well established with an extensive group of users. The SCALE system includes a variety of functional module codes performing fission criticality, transmutation, radiation propagation, and shielding calculations. SCALE has multiple master cross section libraries, including the coupled neutron/gamma libraries used in the present investigation and it has special functional modules such as BONAMI and NITAWL for resonance shielding calculations in problem-specific multigroup cross-section preparation. It has material properties libraries along with code modules to prepare material information in formats suitable for use by other modules. It has a library of utility programs. It also has a library of special FORTRAN subroutines available for use by a newly coded software module.   SCALE includes a set of control module codes, each of which causes subsets of SCALE's functional modules to execute in predefined sequences, while also accomplishing associated auxiliary calculations. SCALE users select any particular sequence by typing its name at the start of their prepared input data file. Some control modules such as SAS1 implement several selectable sequences. One of them is the SAS1X sequence part, which is

similar to portions of the sequence of calculations needed to design optimal radiation shields. The SAS1X module's code was therefore adopted as a starting point. By adding modifications, it was adapted to develop the experimental SAS1XOPT control module used here. This implementation accommodates the sum of both neutron and gamma dose rates in optimal shield design via SCALE's XSDOSE module, which also calculates dose at points in vacuum distant from a shield. Thus, gamma rays released by radiative capture of neutrons are modeled and are part of the shield optimization.

Spherical symmetry reduces the BTE's independent variables from six dimensions to three, i.e., the radial position r; the ordinate μ which is the cosine of the angle from the radius to a surrounding cone of particle directions; and the particle energy E. Eq. (43) is the BTE as simplified by spherical symmetry.

$$\frac{\mu}{r^2}\frac{\partial}{\partial r}\left[r^2\psi(r,\mu,E)\right]+\frac{1}{r^2}\frac{\partial}{\partial\mu}\left[(1-\mu^2)\psi(r,\mu,E)\right]+\sigma(r,E)\psi(r,\mu,E)=q(r,\mu,E) \quad , \qquad (43)$$

where σ is the total cross section and q is the particle source including scattering. SCALE's XSDRN module was the workhorse for numerical experiments in optimal shield design since it rapidly solves the spherically symmetric BTE (or its related adjoint BTE) for any design $\underline{u}(r)$, using iterative discrete ordinates methods. Chapter 3 of Ref. 12 derives these methods and Ref. 29 describes XSDRN's implementation and options. XSDRN's solution is a three-dimensional (3-D) grid of computed $\psi_{i,m,g}$ angular flux (or adjoint angular flux) values, where i indexes into a set of I discrete radial interval midpoints $\{r_i\}_{i=0}^{i=I}$, m indexes into a set of M direction ordinates $\{\mu_m\}_{m=1}^{m=M}$ chosen as roots of the Mth Legendre polynomial, and g indexes into a set of G energy groups $\{E_g\}_{g=1}^{g=G}$. The algorithm's inner loop converges on a space-angle subgrid solution for one energy group, and an outer loop steps through energy groups. Gauss-Seidel updates proceed in a

series of radial sweeps sequenced in ascending order by their constant μ_m values, and updates within each sweep proceed radially in the order closest to particle direction. The inner iteration ends recalculating within-group scattering from the updated angular fluxes in order to update particle sources stored between iterations. Scattering contributions from the converged energy group to other groups are updated before the outer loop advances to the next group. Outer loop iterations between energy groups are avoided where upscattering is insignificant.

It should be noted that in the spherical shield case only a single external location is needed to characterize dose rate, so only one KT multiplier is used. Since this makes the outer loop of Fig. 5 almost trivial the outer loop was incorporated into the experimental SAS1XOPT code along with the inner loop, and a combination of bisection with linear interpolation was used for the single-dimension τ_α to $D_\alpha - D_\alpha^{MAX}$ search. No computational experiments were made with asymmetric shield designs that would require using several KT multipliers.

The chosen model problem was the design of spherical shielding for a family of mobile fission reactor engines powering manned Mars surface vehicles[30] as illustrated in Fig. 6.
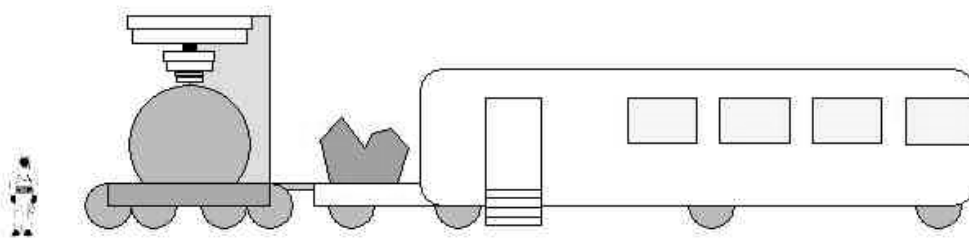

Fig. 6: Shielded Nuclear Reactor Engine with Mars Vehicle

This application is interesting in its own right but also illustrates a situation in which much shielding is essential but its delivery cost is high. Engines developing rated shaft output powers

from 74.6 kW to 7.46 MW [i.e., from 100 to 10,000 horsepower (HP)] were considered which would enable Mars surface missions ranging from excursions in pressurized rover vehicles lasting for long durations to large scale ground excavation, mining, or deep drilling. Fission reactors would use unmoderated highly enriched uranium in uranium nitride plate fuel elements operating at high temperature, similar to the SP100 design.[31] An Open Brayton Cycle (OBC) implemented in three radial flow compressors and turbine stages would transfer reactor heat in lithium-7 coolant through a heat exchanger into compressed Martian air which would then expand through turbines extracting work before exiting carrying the waste heat exhaust. Adequate efficiencies in the low-pressure stages would be obtained by using large diameters.

Radiation shielding requirements for all engines considered were chosen to limit the dose rate (neutron + gamma) at R=6 meters from the reactor center to 13.75 μGy/hr, i.e., to 0.12 Gy per full power Earth-year. The menu of possible shielding materials was provided to the design algorithm, with all costs per unit volume set proportional to material densities. Summaries of minimum-mass design results are given in Tables II and III.

TABLE II
Shielding Materials Menu Choices in Minimum Mass Designs

Shielding Material	Used in Minimum Mass Designs
Tungsten	Used
Boron-10	Used
Beryllium	Not used
Iron	Not used
Polyethylene Plastic	Not used
Lead	Not used
6Lithium Hydride	Used
Uranium-238	Used
Graphite carbon	Not used
7Lithium Hydride	Not used
Boron-10 Carbide	Not used
Water	Not used

TABLE III
Algorithm-Designed Minimum-Mass Spherical Shields Limiting R=6 m Dose Rate
(neutron+gamma) to 13.75 µGy/hr (0.12 Gy/yr)

Design Number	Shaft Power (HP)	Shaft Power (MW)	Reactor Thermal Power (MW)	Reactor +Shield Mass (kg)	Shield Outer Radius (m)
1	100	0.0746	0.310	18,949	1.3314
2	178	0.1328	0.4958	20,220	1.3351
3	316	0.2357	0.8217	21,788	1.3444
4	562	0.4193	1.368	23,676	1.3509
5	1,000	0.7460	2.291	26,154	1.3562
6	1,780	1.328	3.931	29,320	1.3615
7	3,160	2.357	6.879	34,178	1.3677
8	5,620	4.193	12.136	38,219	1.3759
9	10,000	7.46	21.274	45,671	1.3843

The optimal shielding design algorithms assigned materials in layers as thin as 0.5 millimeters, and in some regions different material types alternated to in effect implement material mixtures.

As shown in Table III, spherical shield design #3 is for a 821.7 kWth reactor which would power a turbo engine system using Martian air to develop an estimated 225 kW (i.e., 316 HP) of mechanical shaft power. For the purpose of graphical display the integrated thickness of each shield material type used appears in Fig. 7 above a stacked bar chart showing the actual layout versus radius.

Because water is expected to be abundant on Mars as near-surface buried ice, the 225 kW (i.e., 316 HP) case was rerun with the cost for water as a shielding material reset to $0.01/g, while all other material costs were held at $50/g. The resulting shield design whose material compositions are displayed in Fig. 8 uses 4,800 kg of water plus 17,850 kg of other shielding materials. This increases total shield mass by 862 kg, while reducing the mass delivered from Earth by 3,938 kg, thus reducing total cost.
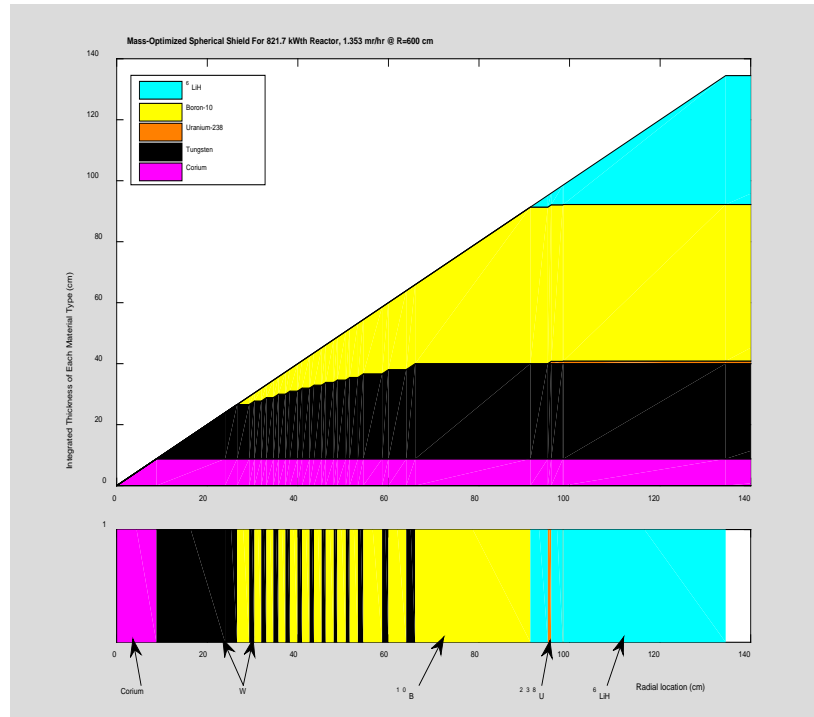
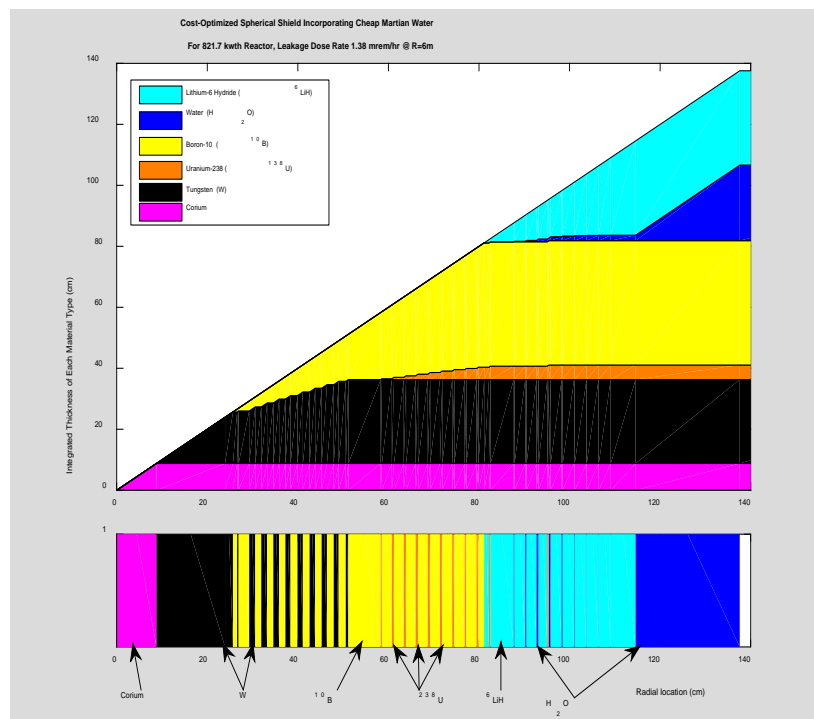Fig. 7: Material Mixes in Mass-Optimized Shield Design

Fig. 8: Material Mixes in Cost-Optimized Shield Design

V. CONCLUSIONS

The future use of 3-D Boltzmann solvers will allow extending optimal design results to shaped shield in which material is not wasted to maintain unneeded spherical symmetry. The largest cost savings of automated optimal design of shielding is expected to be for applications involving nuclear-powered manned space missions where the required shielding attenuation of radiation and material delivery costs are both large. However, less exotic applications such as the shielding of radiation facilities on Earth could also benefit from optimal shielding design, and automation of the design process may reduce engineering costs.

VI. ENDNOTES

While the author is an employee of Princeton Plasma Physics Laboratory, a National Laboratory managed by Princeton University for the Department of Energy, the present work, initiated as part of a Master's thesis, was completed without employer involvement.

VII. REFERENCES

1. N. M. SCHAEFFER, ed., *Reactor Shielding For Nuclear Engineers*, TID-25951, U.S. Atomic Energy Commission, Office of Information Services (1973).

2. W. W. ENGLE, JR., *A Users' Manual for ASOP-ANISN Shield Optimization Program*, CTC-INF-941, Union Carbide Corp. (1969).

3. W. W. ENGLE, JR., *A User' Manual for ANISN, A One-Dimensional Discrete Ordinates Transport Code with Anisotropic Scattering*, K-1693, Union Carbide Corp. (1967).

4. T. A. GABRIEL et al., *The Oak Ridge National Laboratory Strategic Defense Initiative Shield Optimization Program,* Oak Ridge National Laboratory ORNL/TM-10631 (1988).

5. J. KENNETH SHULTIS and RICHARD E. FAW, *Radiation Shielding*, Upper Saddle River, NJ: Prentice Hall PTR (1996).

6. L. S. PONTRYAGIN et al., *The Mathematical Theory of Optimal Processes*, translated by K. N. Trirogoff, New York: Interscience Publishers - a division of John Wiley & Sons (1962).

7. L. S. PONTRYAGIN et al., *The Mathematical Theory of Optimal Processes*, translated by D. E. Brown, New York: The MacMillan Company (1964).

8. ATLE SEIERSTAD and KNUT SYDSAETER, "Sufficient Conditions in Optimal Control Theory," *International Economic Review,* **18,** 2, 367 (1977).

9. O. L. MANGASARIAN, "Sufficient Conditions for the Optimal Control of Nonlinear Systems," *Siam Journal on Control*, **IV**, 139 (1966).

10. K. J. ARROW, "Applications of Control Theory to Economic Growth," in *Mathematics of the Decision Sciences*, eds. G.B. Dantzig and A. F. Veinott, Jr., Providence, RI: American Mathematics Society (1968).

11. ARTHUR E. BRYSON, JR. and YU-CHI HO, *Applied Optimal Control*, Waltham, MA: Ginn and Company – a Xerox Company (1969).

12. E. E. LEWIS and W. F. MILLER, JR., *Computational Methods of Neutron Transport*, La Grange Park, IL: American Nuclear Society, Inc. (1993).

13. H. W. KUHN and A. W. TUCKER, "Nonlinear Programming," in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley, CA: University of California Press, 481 (1951); http://projecteuclid.org/euclid.bsmsp/1200500249.

14. ROBERT D. RICHTMYER and K. W. MORTON, *Difference Methods for Initial-Value Problems,* New York: Interscience Publishers - a division of John Wiley & Sons (1967).

15. GILBERT STRANG and GEORGE J. FIX, *Analysis of the Finite Element Method*, Englewood Cliffs, NJ: Prentice-Hall Inc. (1973).

16. M. I. KRASTANOV, N. K. RIBARSKA and TS. Y. TSACHEV, "A Pontryagin Maximum Principle for Infinite-Dimensional Problems," *SIAM J. Control Optim.*, **49**, 5, 2155 (2011).

17. J. L. LIONS, *Optimal Control of Systems Governed by Partial Differential Equations,* translated by S. K. Mitter, New York: Springer-Verlag (1971).

18. XUNJING LI and JIONGMIN YONG, *Optimal Control Theory for Infinite Dimensional Systems*, Boston: Birkhauser (1995).

19. EDUARDO CASAS and FREDI TROLTZSCH, "First- and Second-Order Optimality Conditions for a Class of Optimal Control Problems With Quasilinear Elliptic Equations," *SIAM J. Control Optim.*, **48,** 2, 688 (2009).

20. M. I. SUMIN, "The First Variation and Pontryagin's Maximum Principle in Optimal Control for Partial Differential Equations," *Computational Mathematics and Mathematical Physics,* **49,** 6, 958 (2009).

21. SABINE PICKENHAIN and MARCUS WAGNER, "Optimal Control Problems with a First Order PDE System – Necessary and Sufficient Optimality Conditions," in *Online Optimization of Large Scale Systems*, M. Grotschel et al. eds., Heidelberg: Springer-Verlag (2001).

22. LAMBERTO CESARI, "Optimization with Partial Differential Equations in Dieudonné-Rashevsky Form and Conjugate Problems," *Arch. Rational Mech. Anal.*, **33**, 23 (1969).

23. ROLF KLOTZLER and SABINE PICKENHAIN, "Pontryagin's Maximum Principle for Multidimensional Control Problems," *International Series in Numerical Mathematics*, **111**, 21, Basel: Birkhauser (1993).

24. M. WAGNER, "Pontryagin's Maximum Principle for Dieudonné-Rashevsky Type Problems Involving Lipschitz Functions," *Optimization: A Journal of Mathematical Programming and Operations Research,* **46**, 2, 165 (1999); doi:10.1080/02331939908844450.

25. MARCUS WAGNER, "Pontryagin's Principle for Dieudonné-Rashevsky Type Problems with Polyconvex Integrands," University of Leipzig, Department of Mathematics (2013); https://www.math.uni-leipzig.de/preprints/p1307.0020.pdf.

26. MIHAI POPESCU, "Existence and Uniqueness of the Optimal Control in Hilbert Spaces for a Class of Linear Systems," *Intelligent Information Management*, **2**, 134 (2010); doi:10.4236/iim.2010.22016 Published Online (http://www.scirp.org/journal/iim).

27. J. TERVO, M. VAUHKONEN and E. BOMAN, "Optimal Control Model for Radiation Therapy Inverse Planning Applying the Boltzmann Transport Equation," *Linear Algebra and its Applications,* **428,** 1230 (2008).

28. J. TERVO and P. KOKKONEN, "On Existence of L^1-Solutions for Coupled Boltzmann Transport Equation and Radiation Therapy Treatment Optimization" arXiv.1406.3228v1 [math.OC] 12 June (2014).

29. *SCALE: A modular code system for performing Standardized Computer Analyses for Licensing Evaluations*, ORNL/TM-2005/39, Version 5, **I-III**, Radiation Safety Information Computational Center at Oak Ridge National Laboratory CCC-725 (2005).

30. ROBERT D. WOOLLEY, "Optimal Design of Mobile Nuclear Reactor Engines to Power Manned Vehicles on Mars", Masters Thesis, The University of Tennessee-Knoxville (2008).

31. G. L. KULCINSKI, "NEEP602 Course Notes (Spring 2000) Nuclear Power in Space," University of Wisconsin-Madison, Fusion Technology Institute (2000); http://fti.neep.wisc.edu/neep602/SPRING00/neep602.html .

Princeton Plasma Physics Laboratory
Office of Reports and Publications

Managed by
Princeton University

under contract with the
U.S. Department of Energy
(DE-AC02-09CH11466)