just to make its arguments explicit.
$\mathcal{D}(\mu,\rho) = \{D_h(X \mid \mu,\rho) : h \in \mathcal{H}^F\}$: all possible full data functions obtained by varying h, but fixing μ,ρ.
$\mathcal{D} = \{D_h(X \mid \mu(F_X),\rho(F_X)) : F_X \in \mathcal{M}^F, h \in \mathcal{H}^F\}$: all possibly full data structure estimating functions obtained by varying h and F_X.
$\mathbf{S^{*F}_{eff}(\cdot \mid F_X)}$: the canonical gradient (also called efficient influence curve) of the pathwise derivative of the parameter $\mu(F_X)$ in the full data model $\mathcal{M}^F$.
$\mathbf{T^{F,\perp,*}_{nuis}(F_X)}$: the set of all gradients of the pathwise derivative at F_X of the parameter $\mu(F_X)$ in the full data model $\mathcal{M}^F$ whose components span $T^{F,\perp}_{nuis}(F_X)$.
$\mathbf{D_{h_{eff}}(\cdot \mid \mu(F_X),\rho(F_X))} = S^{*F}_{eff}(\cdot \mid F_X)$: that is, h_{eff} indexes the optimal estimating function in the *full data structure model.* Here $h_{eff} = h_{eff}(F_X)$ depends on F_X. Off course, one still obtains an optimal estimating function by putting a $k \times k$ fixed matrix in front of S^{*F}_{eff}.
$\mathcal{F}(\mathbf{t})$: a predictable observed subject-specific history up to time t, typically representing all observed data up to time point t on a subject.
A: a time-dependent possibly multivariate process $A(t) = (A_1(t),\ldots,A_k(t))$ whose components describe specific censoring (e.g., treatment) actions at time t. Here A represents the censoring variable C for the observed data structure. Typically, $A_j(t)$, $j = 1,\ldots,k$, are counting processes.
$\mathcal{A}$: the support of the marginal distribution of A.
$\alpha(\mathbf{t}) = E(dA(t) \mid \mathcal{F}(t))$: the intensity (possibly discrete, $\alpha(t) = P(dA(t) = 1 \mid \mathcal{F}(t))$ at given grid points) of counting process $A(t)$ w.r.t. the history $\mathcal{F}(t)$.
$\mathbf{Y = (A, X_A)}$: a particular type of observed censored data, where for the full data we have $X = (X_a : a \in \mathcal{A})$, and A tells us what component of X we observe. For example, $A(t)$ can be the indicator $I(C \le t)$ of being right-censored by a dropout time C. If the full data model is a causal model and there is no censoring, then $A(t)$ is the treatment that the subject receives at time t. If the observed data structure includes both treatment assignment and censoring, then $A(t)$ is the multivariate process describing the treatment actions and censoring actions assigned to the subject at time t.
$\mathbf{L^2_0(P_{F_X,G})}$: Hilbert space of functions $V(Y)$ with $E_{P_{F_X,G}}V(Y) = 0$ with inner-product $\langle h,g\rangle_{P_{F_X,G}} = E_{P_{F_X,G}}h(Y)g(Y)$ and corresponding norm $\| h \|_{P_{F_X,G}} = \sqrt{E_{P_{F_X,G}}h^2(Y)}$.
$\mathbf{T(P_{F_X,G})} \subset L^2_0(P_{F_X,G})$, $\mathbf{T_{nuis}(P_{F_X,G})} \subset L^2_0(P_{F_X,G})$, $\mathbf{T^{\perp}_{nuis}(P_{F_X,G})} \subset L^2_0(P_{F_X,G})$ are the observed data tangent space, observed data nuisance tangent space, and the orthogonal complement of the observed data nuisance tangent space at $P_{F_X,G}$, respectively, in model $\mathcal{M}(CAR)$ (or, if made explicit in $\mathcal{M}(\mathcal{G})$), where μ is the parameter of interest.
$\mathbf{T_{CAR}(P_{F_X,G})} = \{V(Y) : E_G(V(Y) \mid X) = 0\} \subset L^2_0(P_{F_X,G})$: the nuisance tangent space of G in model $\mathcal{M}(CAR)$.

$\mathbf{T_2(P_{F_X,G})} \subset T_{CAR}(P_{F_X,G})$ or $\mathbf{T_G(P_{F_X,G})} \subset T_{CAR}(P_{F_X,G})$: the nuisance tangent space of G in the observed data model $\mathcal{M}(\mathcal{G})$.
$D \to \mathbf{IC_0(Y \mid Q_0, G, D)}$, $D \to \mathbf{IC(Y \mid F_X, G, D)}$, $D \to \mathbf{IC(Y \mid Q, G, D)}$: mapping from a full data function into an observed data function indexed by nuisance parameters $Q_0(F_X,G), G$, F_X, G or $Q(F_X,G), G$. $IC_0(Y \mid Q_0, G, D)$ stands for an initial mapping and $IC(Y \mid F_X, G, D)$ and $IC(Y \mid Q(F_X,G), G, D)$ for the optimal mapping orthogonalized w.r.t. T_{CAR} or a mapping orthogonalized w.r.t. a subspace of T_{CAR}. In many cases, it is not convenient to parametrize IC in terms of F_X, G, but instead parametrize it by a parameter $Q = Q(F_X,G)$ and G. We note that the dependence of these functions on F_X and G is only through the F_X-part of the density of Y and the conditional distribution of Y, given X, respectively.
The mapping IC_0 satisfies for each $P_{F_X,G} \in \mathcal{M}(\mathcal{G})$: for a non empty set of full data functions $\mathcal{D}(\rho_1(F_X), G)$, we have

$$E_G(IC_0(Y \mid Q, G, D) \mid X) = D(X) \ F_X\text{-a.e. for all } Q \in \mathcal{Q}_0. \quad (1)$$

For IC, we have the additional property at each $P_{F_X,G} \in \mathcal{M}(CAR)$:

$$\begin{aligned} IC(Y \mid Q(F_X,G), G, D) &= IC_0(Y \mid Q_0(F_X,G), G, D) \\ &\quad - \Pi_{F_X,G}(IC_0(Y \mid Q_0(F_X,G), G, D) \mid T_{CAR}), \end{aligned}$$

or the projection term can be a projection on a subspace of T_{CAR}. Here $\Pi(\cdot \mid T_{CAR})$ denotes the projection operator in the Hilbert space $L^2_0(P_{F_X,G})$ with inner product $\langle f,g\rangle_{P_{F_X,G}} = E_{P_{F_X,G}}f(Y)g(Y)$.
$\mathcal{D}(\rho_1(\mathbf{F_X}), \mathbf{G})$: the set of full data functions in $\mathcal{D}$ for which (1) holds. Thus, these are the full data structure functions that are mapped by IC_0 into unbiased observed data estimating functions. By making the appropriate assumption on the censoring mechanism, one will have that $\mathcal{D}(\rho_1(F_X), G) = \mathcal{D}$, but one can also decide to make this membership requirement $D_h(\cdot \mid \mu(F_X), \rho(F_X)) \in \mathcal{D}(\rho_1(F_X), G)$ a nuisance parameter of the full data structure estimating function: see next entry.
$\mathbf{D_h(\cdot \mid \mu(F_X), \rho(F_X, G))}$, $h \in \mathcal{H}^F$: these are full data structure estimating functions satisfying $D_h(\cdot \mid \mu(F_X), \rho(F_X,G)) \in \mathcal{D}(\rho_1(F_X), G))$ for all $h \in \mathcal{H}^F$. Formally, they are defined in terms of initially defined full data estimating functions D_h as

$$D^r_h(\cdot \mid \mu, \rho, \rho_1, G) \equiv D_{\Pi(h \mid \mathcal{H}^F(\mu,\rho,\rho_1,G))}(\cdot \mid \mu, \rho),$$

where $\mathcal{H}^F(\mu,\rho,\rho_1,G) \subset \mathcal{H}^F$ are the indexes that guarantee that $E_G(IC_0(Y \mid Q_0, G, D_h(\cdot \mid \mu,\rho)) \mid X) = D_h(X \mid \mu,\rho)$ F_X-a.e and $\Pi(\mid \mathcal{H}^F(\mu,\rho,\rho_1,G))$ is a mapping from $\mathcal{H}^F$ into $\mathcal{H}^F(\mu,\rho,\rho_1,G)$ that is the identity mapping on $\mathcal{H}^F(\mu,\rho,\rho_1,G)$. Thus, if $D_h(\cdot \mid \mu(F_X),\rho(F_X)) \in \mathcal{D}(\rho_1(F_X), G)$ for all $P_{F_X,G} \in \mathcal{M}(\mathcal{G})$, then $D^r_h = D_h$. For notational convenience, we denote $D^r_h(\cdot \mid \mu,\rho,\rho_1,G)$ with $D_h(\cdot \mid \mu,\rho)$ again, but where ρ