Open Journal of
Mathematical Sciences

PSR Press

Article

A formal proof of Castigliano Theorem and a related generalization including a non-linear case

Fabio Silva Botelho

Department of Mathematics, Federal University of Santa Catarina, UFSC, Florianópolis, SC - Brazil; fabio.botelho@ufsc.br

Received: 29 Jan 2026; Revised: 15 Feb 2026; Accepted: 19 Feb 2026; Published: 29 Mar 2026

Abstract: This article develops a formal proof of Castilgiano Theorem in an elasticity theory context. The results are based on standard tools of applied functional analysis and calculus of variations. It is worth mentioning such results here presented may be easily extended to a non-linear elasticity context. Finally, in the last section we present a numerical example in order to illustrate the results applicability.

Keywords: Castigliano theorem, generalized Castigliano theorem, virtual work principle, elasticity theory, numerical example

MSC: 74B05.

---

1. Introduction

In this article version we present some corrections and improvements concerning the previous version [1]. In the next section we present the mathematical formalism of a result in elasticity theory known as the Castigliano's Theorem.

We also present a generalization of such a theorem and its connection with the principle of virtual work in a elasticity theory context.

Furthermore, in this article version we also include a new extension to a non-linear elasticity case.

The results are obtained through an application of basic tools of functional analysis and calculus variations to a solid mechanics theory.

Our main reference in solid mechanics is [2]. Related results have been presented in [3–7].

For the Sobolev spaces involved, we would cite [8].

Remark 1. In this text we have adopted the Einstein convention of summing up repeated indices, unless otherwise indicated.

2. A formal proof of Castigliano Theorem

Let $\Omega = [0, L] \subset \mathbb{R}$ represents the straight axis of an Euler-Bernoulli beam of length $L > 0$ and rectangular cross section $b \times h$.

For $E > 0$, $I = bh^3/12$, $V = W_0^{2,2}(\Omega)$ and for $x_1, \cdots, x_N \in (0, L)$, we define the functional $J : V \to \mathbb{R}$ by

$$J(w) = E_{in}(w) - \langle w, f_0 \rangle_{L^2} - \sum_{j=1}^{N} w(x_j)(P_0)_j,$$

where $w \in V$ represents the vertical displacements field resulting from the actions of a vertical load $f_0 \in C^1(\Omega)$ and punctual loads

$$(P_0)_j \in \mathbb{R}, \ \ j \in \{1, \cdots, N\}$$

for some $N \in \mathbb{N}$.

Here we have denoted

$$P_0 = \{(P_0)_j\} \in \mathbb{R}^N.$$

and
$$E_{in}(w) = \frac{EI}{2}\int_\Omega (w'')^2\,dx.$$

Observe that the variation of J in w give us the following Euler-Lagrange equation

$$EIw'''' - f_0 - \sum_{j=1}^N (P_0)_j \delta(x - x_j) = \mathbf{0}, \text{ in } \Omega, \tag{1}$$

in a distributional sense. Here $\delta(x - x_j)$ denotes a standard Dirac delta in a distributional sense.

Symbolically such an equation stands for

$$\frac{\partial J(w)}{\partial w} = \mathbf{0},$$

so that

$$\frac{\partial\left(E_{in} - \langle w, f_0\rangle_{L^2} - \sum_{j=1}^N w(x_j)(P_0)_j\right)}{\partial w} = \mathbf{0}. \tag{2}$$

Let $H_1(x)$ and $(H_2)_j(x)$ be functions such that, classically,

$$H_1''''(x) = f_0(x),$$

and

$$(H_2)_j'''(x) = u_j(x),\ \forall j \in \{1, \cdots, N\}, \text{ in } \Omega,$$

where

$$u_j(x) = \begin{cases} 0, & \text{if } 0 \leq x < x_j, \\ 1, & \text{if } x_j \leq x \leq L. \end{cases} \tag{3}$$

Let $\varphi \in C_c^\infty((0, L))$.

Observe that

$$\begin{aligned} -\int_0^L u_j(x)\varphi'(x)\,dx &= -\int_{x_j}^L \varphi'(x)\,dx \\ &= -\varphi(L) + \varphi(x_j) \\ &= \varphi(x_j). \end{aligned} \tag{4}$$

From such results, we may infer that

$$u_j'(x) = \delta(x - x_j),$$

in a distributional sense, so that

$$(H_2)_j''''(x) = \delta(x - x_j),$$

in a distributional sense.

Observe also that for a solution w of (1), we have

$$w(x) = \frac{H_1(x)}{EI} + \sum_{j=1}^N \frac{(P_0)_j}{EI}(H_2)_j(x) + c_1 + c_2 x + c_3 x^2 + c_4 x^3,$$

where c_1, c_2, c_3 and $c_4 \in \mathbb{R}$ must be such that

$$w(0) = w'(0) = w(1) = w'(1) = 0.$$

For a fixed f_0, here we define

$$F_1(w,P)=\frac{1}{2}E_{in}(w)-\langle w,f_0\rangle_{L^2}-\sum_{j=1}^{N}w(x_j)P_j,$$

and

$$F(w,P)=\frac{\partial F_1(w,P)}{\partial w}.$$

Hence a solution of equation

$$F(w,P)=EIw''''-f_0-\sum_{j=1}^{N}(P)_j\delta(x-x_j)=\mathbf{0},\text{ in }\Omega,$$

is such that

$$w=w(x,P)=H_1(x)+\sum_{j=1}^{N}(P)_j(H_2)_j(x)+c_1(P)+c_2(P)x+c_3(P)x^2+c_4(P)x^3,$$

where $c_1(P),c_2(P),c_3(P)$ and $c_4(P)\in\mathbb{R}$ must be such that

$$w(0)=w'(0)=w(1)=w'(1)=0,$$

so that $c_j(P)$ is linear in P, $\forall j\in\{1,\cdots,N\}$.

We recall that we have

$$EIw''''-f_0-\sum_{j=1}^{N}P_j\delta(x-x_j)=\mathbf{0},\text{ in }\Omega, \tag{5}$$

in a distributional sense.

Thus,

$$\int_0^L EIw''\varphi''\,dx-\int_0^L f\,\varphi\,dx-\sum_{j=1}^{N}P_j\varphi(x_j)=\mathbf{0}, \tag{6}$$

$\forall\varphi\in W_0^2(\Omega)$.

In particular for $\varphi=w$, we obtain

$$\int_0^L EI(w'')^2\,dx-\int_0^L f_0\,w\,dx-\sum_{j=1}^{N}(P)_jw(x_j)=\mathbf{0}, \tag{7}$$

Observe also that, in such a case, we have

$$\begin{aligned}H_1(w(P),P)&\equiv 2E_{in}(w(P))-\langle w(P),f_0\rangle_{L^2}-\sum_{j=1}^{N}w(x_j,P)P_j\\&=0,\ \forall P\in\mathbb{R}^N,\end{aligned} \tag{8}$$

and thus,

$$\frac{d}{dP_s}\left(H_1(w(P),P)\right)=0.$$

so that,

$$2\frac{dE_{in}}{dP_s}-\frac{d}{dP_s}\left(\langle w(P),f\rangle_{L^2}+\sum_{j=1}^{N}w(x_j,P)P_j\right)=0. \tag{9}$$

Moreover, from its expression $w(P)$ is Fréchet differentiable and the derivative

$$\frac{\partial w(x,P)}{\partial P_s}$$

is well defined so that from a standard Chain Rule Theorem for Fréchet differentiable functions, we obtain

$$\frac{dE_{in}}{dP_s} = \frac{\partial E_{in}}{\partial w}\frac{\partial w}{\partial P_s}.$$

Hence, from this and (9), we obtain

$$\frac{dE_{in}}{dP_s} + \frac{d}{dP_s}\left(E_{in} - \langle w(f), f_0\rangle_{L^2} - \sum_{j=1}^{N} w(x_j, P)P_j\right) = 0,$$

so that

$$\frac{dE_{in}}{dP_s} + \left(\frac{\partial(E_{in} - \langle w, f_0\rangle_{L^2} - \sum_{j=1}^{N} w(x_j)P_j)}{\partial w}\frac{\partial w}{\partial P_s}\right) - \frac{\partial}{\partial P_s}\left(\langle w, f_0\rangle_{L^2} + \sum_{j=1}^{N} w(x_j)P_j\right) = 0. \tag{10}$$

From this, recalling that

$$\frac{\partial(E_{in} - \langle w, f_0\rangle_{L^2} - \sum_{j=1}^{N} w(x_j)P_j)}{\partial w} = 0,$$

we obtain

$$\frac{dE_{in}}{dP_s} - w(x_s, P) = 0,$$

so that recalling that

$$w_0(x_s) = w(x_s, P_0),$$

we have obtained

$$w_0(x_s) = \left[\frac{dE_{in}}{dP_s}\right]_{P=P_0} = \left[\frac{d}{dP_s}\left(\frac{1}{2}\int_\Omega EI(w''(P))^2\,dx\right)\right]_{P=P_0},$$

$\forall s \in \{1, \cdots, N\}$.

With such results in mind, we have proven the following theorem.

Theorem 1 (Castigliano)**.** *Considering the notations and definitions in this section, we have*

$$w_0(x_s) = \left[\frac{dE_{in}}{dP_s}\right]_{P=P_0} = \left[\frac{d}{dP_s}\left(\frac{1}{2}\int_\Omega EI(w''(P))^2\,dx\right)\right]_{P=P_0},$$

$\forall s \in \{1, \cdots, N\}$.

2.1. The virtual work principle

Considering the definitions, results and statements of the previous section, we may also easily prove the following theorem.

Theorem 2 (The virtual work principle)**.** *Let $x_l \in \Omega^0 = (0, L)$ such that $x_l \neq x_j \in \Omega^0, \forall j \in \{1, \cdots, N\}$.*

For a virtual constant load $\hat{P} \in \mathbb{R}$ on x_l at the direction of $w(x_l)$, define now $J : V \to \mathbb{R}$ where

$$J(u) = E_{in} - \langle w, f\rangle_{L^2} - \sum_{j=1}^{N} w(x_j)P_j - \hat{P}w(x_l).$$

Under such hypotheses, we have

$$w(x_l) = \left(\frac{d\ E_{in}(w(\hat{P}))}{d\hat{P}}\right)_{\hat{P}=0},$$

$\forall x_l \in \Omega$ *such that* $x_l \neq x_j$.

Proof. The proof is exactly the same as in the Castigliano Theorem in the previous section except by setting the virtual load $\hat{P} = 0$ in the end of this calculation and will not be repeated. □

3. A generalization of Castigliano Theorem

In this section we present a generalization of Castigliano Theorem in a linear elasticity context.

Let $\Omega \subset \mathbb{R}^3$ be an open, bounded and connected set with a regular (Lipschitzian) boundary denoted by $\partial\Omega$.

In a context of linear elasticity, consider the functional $J : V \to \mathbb{R}$ where

$$J(u) = E_{in}(u) - \langle u_i, f_i \rangle_{L^2},$$

$u = (u_1, u_2, u_3) \in W_0^{1,2}(\Omega; \mathbb{R}^3) \equiv V$, $f = (f_1, f_2, f_3) \in L^2(\Omega; \mathbb{R}^3)$, $Y = Y^* = L^2(\Omega; \mathbb{R}^3)$,

Here we have denoted

$$\begin{aligned} E_{in}(u) =& \frac{1}{2} \int_\Omega H_{ijkl} e_{ij}(u) e_{kl}(u)\, dx, \\ e_{ij}(u) =& \frac{1}{2} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right). \end{aligned}$$

Moreover H_{ijkl} is a fourth order positive definite and constant tensor.

Observe that the variation of J in u_i give us the following Euler-Lagrange equation

$$-(H_{ijkl} e_{kl}(u))_{,j} - f_i = \mathbf{0}, \text{ in } \Omega. \tag{11}$$

Symbolically such a system stands for

$$\frac{\partial J(u)}{\partial u_i} = \mathbf{0}, \forall i \in \{1, 2, 3\},$$

so that

$$\frac{\partial (E_{in} - \langle u_i, f_i \rangle_{L^2})}{\partial u_l} = \mathbf{0}, \ \forall l \in \{1, 2, 3\}. \tag{12}$$

We denote a solution $u \in V$ of (11) by $u = u(f)$, so that multiplying the concerning extremal equation by u_i and integrating by parts, we get

$$\begin{aligned} H_1(u(f), f) &= 2E_{in}(u(f)) - \langle u_i(f), f_i \rangle_{L^2} \\ &= 0, \ \forall f \in Y^*. \end{aligned} \tag{13}$$

Therefore

$$\frac{d}{df_s} \left(H_1(u(f), f) \right) = 0,$$

so that

$$2\frac{dE_{in}}{df_s} - \frac{d}{df_s} \left(\langle u_i(f), f_i \rangle_{L^2} \right) = 0, \tag{14}$$

where we recall that, assuming the hypotheses of the Implicit Function Theorem in Banach Spaces as it may be found in [9,10] or at page 346 of [11] (please see the Appendix A for details) at a point

$$(u_0, f_0) = (u(f_0), f_0) \in V \times Y,$$

and from the Chain Rule Theorem for Fréchet differentiable functionals, we have

$$\frac{dE_{in}}{df_s} = \frac{\partial E_{in}}{\partial u_k} \frac{\partial u_k}{\partial f_s},$$

in a neighborhood of (u_0, f_0).

Hence, from this and (14), we obtain

$$\frac{dE_{in}}{df_s} + \frac{d}{df_s}\left(E_{in} - \langle u_i(f), f_i\rangle_{L^2}\right) = 0,$$

so that

$$\frac{dE_{in}}{df_s} + \left(\frac{\partial(E_{in} - \langle u_i, f_i\rangle_{L^2})}{\partial u_k}\,\frac{\partial u_k}{\partial f_s}\right) - \frac{\partial}{\partial f_s}\left(\langle u_i, f_i\rangle_{L^2}\right) = 0. \tag{15}$$

From this, recalling that

$$\frac{\partial(E_{in} - \langle u_i, f_i\rangle_{L^2})}{\partial u_k} = 0,\ \forall k \in \{1,2,3\}$$

we obtain

$$\frac{dE_{in}}{df_s} - u_s = 0,$$

so that

$$u_s = \frac{dE_{in}}{df_s} = \frac{d}{df_s}\left(\frac{1}{2}\int_\Omega H_{ijkl}e_{ij}(u(f))e_{kl}(u(f))\,dx\right),$$

$\forall l \in \{1,2,3\}$.

With such results in mind, we have proven the following theorem.

Theorem 3 (Castigliano). *Considering the notations and definitions in this section and assuming the hypotheses of the Implicit Function Theorem in Banach Spaces at a point*

$$(u_0, f_0) = (u(f_0), f_0) \in V \times Y,$$

we have

$$u_s = \frac{dE_{in}}{df_s} = \frac{d}{df_s}\left(\frac{1}{2}\int_\Omega H_{ijkl}e_{ij}(u(f))e_{kl}(u(f))\,dx\right),$$

$\forall s \in \{1,2,3\}$, *in a neighborhood of* (u_0, f_0), *so that in particular*

$$(u_0)_s = \left[\frac{dE_{in}}{df_s}\right]_{f=f_0} = \left[\frac{d}{df_s}\left(\frac{1}{2}\int_\Omega H_{ijkl}e_{ij}(u(f))e_{kl}(u(f))\,dx\right)\right]_{f=f_0},$$

$\forall s \in \{1,2,3\}$.

4. The Castigliano Theorem for a non-linear elasticity case

Let $\Omega \subset \mathbb{R}^3$ be an open, bounded and connected set with a regular (Lipschitzian) boundary denoted by $\partial\Omega$.

In a context of non-linear elasticity, consider the functional $J : V \to \mathbb{R}$ where

$$J(u) = E_{in} - \langle u_i, f_i\rangle_{L^2},$$

$u = (u_1, u_2, u_3) \in W_0^{1,2}(\Omega;\mathbb{R}^3) \equiv V$, $f = (f_1, f_2, f_3) \in L^2(\Omega;\mathbb{R}^3)$ and $Y = Y^* = L^2(\Omega;\mathbb{R}^3)$.

Here we have denoted

$$E_{in} = \frac{1}{2}\int_\Omega H_{ijkl}e_{ij}(u)e_{kl}(u)\,dx,$$

$$e_{ij}(u) = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} + u_{m,i}\,u_{m,j}\right).$$

Moreover H_{ijkl} is a fourth order positive definite and constant tensor.

Observe that the variation of J in u_i give us the following Euler-Lagrange equation

$$-(H_{ijkl}e_{kl}(u))_{,j} - (H_{mjkl}e_{kl}(u)u_{i,m})_{,j} - f_i) = \mathbf{0}, \text{ in } \Omega. \tag{16}$$

Symbolically such a system stands for

$$\frac{\partial J(u)}{\partial u_i} = \mathbf{0}, \forall i \in \{1,2,3\},$$

so that

$$\frac{\partial(E_{in} - \langle u_i, f_i \rangle_{L^2})}{\partial u_l} = \mathbf{0}, \ \forall l \in \{1,2,3\}. \tag{17}$$

We denote $u \in V$ solution of (16) by $u = u(f)$, so that multiplying the concerning extremal equation by u_i and integrating by parts, we get

$$\begin{aligned} H_1(u(f), f) &= 2E_{in}(u(f)) + \frac{1}{2}\left\langle \sigma_{ij}(u(f))(u_m(f))_{,i}(u_m(f))_{,j}\right\rangle_{L^2} - \langle u_i(f), f_i\rangle_{L^2} \\ &= 0, \ \forall f \in Y^*, \end{aligned} \tag{18}$$

where

$$\sigma_{ij}(u) = H_{ijkl}e_{kl}(u), \ \forall i,j \in \{1,2,3\}.$$

Therefore

$$\frac{d}{df_s}\left(H_1(u(f), f)\right) = 0,$$

so that

$$2\frac{dE_{in}}{df_s} + \frac{d}{df_s}\left(\frac{1}{2}\left\langle \sigma_{ij}(u(f))(u_m(f))_{,i}(u_m(f))_{,j}\right\rangle_{L^2}\right) - \frac{d}{df_s}\left(\langle u_i(f), f_i\rangle_{L^2}\right) = 0, \tag{19}$$

where we recall that, assuming the hypotheses of the Implicit Function Theorem in Banach Spaces at a point

$$(u_0, f_0) = (u(f_0), f_0) \in V \times Y,$$

similarly as indicated in §3 and from the Chain Rule Theorem for Fréchet differentiable functions, we have

$$\frac{dE_{in}}{df_s} = \frac{\partial E_{in}}{\partial u_k}\frac{\partial u_k}{\partial f_s}.$$

Hence, from this and (19), we obtain

$$\frac{dE_{in}}{df_s} + \frac{d}{df_s}\left(\frac{1}{2}\left\langle \sigma_{ij}(u(f))(u_m(f))_{,i}(u_m(f))_{,j}\right\rangle_{L^2}\right) + \frac{d}{df_s}\left(E_{in} - \langle u_i(f), f_i\rangle_{L^2}\right) = 0, \tag{20}$$

so that

$$\frac{dE_{in}}{df_s} + \frac{d}{df_s}\left(\frac{1}{2}\left\langle \sigma_{ij}(u(f))(u_m(f))_{,i}(u_m(f))_{,j}\right\rangle_{L^2}\right) + \left(\frac{\partial(E_{in} - \langle u_i, f_i\rangle_{L^2})}{\partial u_k}\frac{\partial u_k}{\partial f_s}\right) - \frac{\partial}{\partial f_s}\left(\langle u_i, f_i\rangle_{L^2}\right) = 0. \tag{21}$$

From this, recalling that

$$\frac{\partial(E_{in} - \langle u_i, f_i\rangle_{L^2})}{\partial u_k} = 0, \ \forall k \in \{1,2,3\},$$

we obtain

$$\frac{dE_{in}}{df_s} + \frac{d}{df_s}\left(\frac{1}{2}\left\langle \sigma_{ij}(u(f))(u_m(f))_{,i}(u_m(f))_{,j}\right\rangle_{L^2}\right) - u_s = 0,$$

so that

$$\begin{aligned} u_s &= \frac{dE_{in}}{df_s} + \frac{d}{df_s}\left(\frac{1}{2}\left\langle \sigma_{ij}(u(f))(u_m(f))_{,i}(u_m(f))_{,j}\right\rangle_{L^2}\right) \\ &= \frac{d}{df_s}\left(\frac{1}{2}\int_\Omega H_{ijkl}e_{ij}(u(f,))e_{kl}(u(f))\,dx\right) + \frac{d}{df_s}\left(\frac{1}{2}\left\langle \sigma_{ij}(u(f))(u_m(f))_{,i}(u_m(f))_{,j}\right\rangle_{L^2}\right), \end{aligned} \tag{22}$$

$\forall s \in \{1,2,3\}$.

With such results in mind, we have proven the following theorem.

Theorem 4. *Considering the notations and definitions in this section, assuming the hypotheses of the Implicit Function Theorem in Banach Spaces at a point*

$$(u_0, f_0) = (u(f_0), f_0) \in V \times Y,$$

in particular considering

$$E_{in} = \frac{1}{2}\int_\Omega H_{ijkl}e_{ij}(u)e_{kl}(u)\,dx,$$

and

$$e_{ij}(u) = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} + u_{m,i}u_{m,j}\right),$$

we have

$$\begin{aligned} u_s &= \frac{dE_{in}}{df_s} + \frac{d}{df_s}\left(\frac{1}{2}\left\langle \sigma_{ij}(u(f))(u_m(f))_{,i}(u_m(f))_{,j}\right\rangle_{L^2}\right) \\ &= \frac{d}{df_s}\left(\frac{1}{2}\int_\Omega H_{ijkl}e_{ij}(u(f))e_{kl}(u(f))\,dx\right) + \frac{d}{df_s}\left(\frac{1}{2}\left\langle \sigma_{ij}(u(f))(u_m(f))_{,i}(u_m(f))_{,j}\right\rangle_{L^2}\right), \end{aligned} \tag{23}$$

$\forall s \in \{1,2,3\}$, *in a neighborhood of* (u_0, f_0), *so that in particular,*

$$\begin{aligned} (u_0)_s &= \left[\frac{dE_{in}}{df_s} + \frac{d}{df_s}\left(\frac{1}{2}\left\langle \sigma_{ij}(u(f))(u_m(f))_{,i}(u_m(f))_{,j}\right\rangle_{L^2}\right)\right]_{f=f_0} \\ &= \left[\frac{d}{df_s}\left(\frac{1}{2}\int_\Omega H_{ijkl}e_{ij}(u(f))e_{kl}(u(f))\,dx\right) + \frac{d}{df_s}\left(\frac{1}{2}\left\langle \sigma_{ij}(u(f))(u_m(f))_{,i}(u_m(f))_{,j}\right\rangle_{L^2}\right)\right]_{f=f_0}, \end{aligned} \tag{24}$$

$\forall s \in \{1,2,3\}$.

4.1. A numerical example related to the Castigliano Theorem

Let $\Omega = [0,1] \subset \mathbb{R}$ be the axis of a straight beam with a rectangular cross section of dimensions $b \times h$, where units in this subsections refer to the international system.

Let

$$I = \frac{bh^3}{12},$$

and denote by $E > 0$ the Young modulus for a steel beam.

Assume such a beam is subject to a vertical load $P > 0$ uniformly distributed on Ω.

Assume also the beam is clamped at $x = 0$ and simply supported at $x = 1$.

Denoting by $w \in V = W^{2,2}(\Omega)$ the vertical field of displacements results from the action of P, the related boundary conditions are given by

$$w(0) = w(1) = 0,\ w_{,x}(0) = 0,\ w_{,xx}(1) = 0.$$

The total beam energy is defined by $J : V \to \mathbb{R}$, where

$$J(w) = \frac{EI}{2}\int_\Omega w_{,xx}^2\,dx - \int_\Omega Pw\,dx.$$

In order to apply the results of the previous section, we free the rotations of the beam at $x = 0$, considering a moment load M on $x = 0$ with general work

$$Mw_{,x}(0).$$

Hence, we define

$$V_1 = \{w \in W^{2,2}(\Omega)\ :\ w(0) = w(1) = w_{,xx}(1) = 0\},$$

and define $J_1 : V_1 \to \mathbb{R}$, by

$$J_1(w) = \frac{EI}{2}\int_\Omega w_{,xx}^2\,dx - \int_\Omega Pw\,dx - M\,w_{,x}(0).$$

We emphasize again, through the methods of the previous section, we intend to obtain the value of M which corresponds to

$$w_{,x}(0) = 0.$$

Let $\varphi \in W^{2,2}$. The variation

$$\delta J_1(w;\varphi),$$

stands for

$$\delta J_1(w;\varphi) = EI\int_\Omega w_{,xx}\varphi_{,xx}\,dx - \int_\Omega P\varphi\,dx - M\varphi_{,x}(0).$$

Here assuming, w is smooth enough, integrating by parts and recalling that $\varphi \in V_1$, we obtain

$$\delta J_1(w;\varphi) = EI\int_\Omega w_{,xxxx}\varphi\,dx - EIw_{,xx}(0)\varphi_{,x}(0) - \int_\Omega P\varphi\,dx - M\varphi_{,x}(0),$$

so that the extremal condition

$$\delta J_1(w;\varphi) = 0,\ \forall \varphi \in V_1,$$

provide us the following natural boundary condition

$$M = -EIw_{,xx}(0),$$

and the equation

$$EIw_{,xxxx} - P = 0,\ \text{in } \Omega.$$

Here we recall the remaining essential boundary conditions,

$$w(0) = w(1) = w_{,xx}(1) = 0.$$

A particular solution of such an equation

$$EIw_{,xxxx} - P = 0,\ \text{in } \Omega,$$

stands for

$$w_p(x) = ax^4,$$

where

$$a = \frac{P}{4!\ EI}.$$

The concerning general solution stands for

$$w(x) = w_p(x) + bx^3 + cx^2 + dx + e.$$

The boundary condition $w(0) = 0$ implies that $e = 0$.

The boundary condition

$$M = -EIw_{,xx}(0),$$

stands for

$$M = -EI(2c),$$

so that

$$c = -\frac{M}{2EI}.$$

Moreover, from

$$w_{,xx}(1) = 0,$$

we obtain

$$12a + 6b + 2c = 0.$$

From

$$w(1) = 0,$$

we have

$$a + b + c + d = 0.$$

From these last two equations, we obtain

$$b = -\frac{c}{3} - 2a \equiv b(c),$$

and

$$d = a - \frac{2}{3}c.$$

From the Castigliano Theorem, $M \in \mathbb{R}$ must be such that

$$w_x(0) = \frac{dE_{in}}{dM} = 0,$$

where

$$E_{in} = \frac{EI}{2}\int_\Omega w_{,xx}^2\,dx.$$

Observe that

$$\frac{dE_{in}}{dM} = \frac{dE_{in}}{dc}\frac{dc}{dM} = 0,$$

so that in fact, it suffices to obtain,

$$\frac{dE_{in}}{dc} = 0.$$

Moreover, observe that

$$\begin{aligned}
E_{in} &= \frac{EI}{2}\int_\Omega w_{,xx}^2\,dx \\
&= \frac{EI}{2}\int_0^1 (12ax^2 + 6b(c)x + 2c)^2\,dx \\
&= \frac{EI}{2}\int_0^1 (144a^2x^4 + 36b(c)^2x^2 + 4c^2 + 144ab(c)x^3 + 48acx^2 + 24xb(c)c)\,dx \\
&= \frac{EI}{2}\left[\frac{144a^2x^5}{5} + 36\frac{b(c)^2x^3}{3} + 4c^2x + \frac{144ab(c)x^4}{4} + \frac{48acx^3}{3} + \frac{24b(c)cx^2}{2}\right]_0^1 \\
&= \frac{EI}{2}\left(\frac{144a^2}{5} + \frac{36b(c)^2}{3} + 4c^2 + \frac{144ab(c)}{4} + \frac{48ac}{3} + \frac{24cb(c)}{2}\right).
\end{aligned} \tag{25}$$

Hence, we must have

$$\begin{aligned}
\frac{dE_{in}}{dc} &= \frac{\partial E_{in}}{\partial b}\frac{db}{dc} + \frac{\partial E_{in}}{\partial c} \\
&= 0
\end{aligned} \tag{26}$$

so that,

$$12(2)b(c)\frac{db(c)}{dc}+8c+36a\frac{db(c)}{dc}+16a+12b(c)+12c\frac{db(c)}{dc}=0. \tag{27}$$

Recalling that

$$\frac{db(c)}{dc}=-1/3,$$

we have got

$$-8b(c)+8c-12a+16a+12b(c)-4c=0.$$

From such a result and recalling that

$$b(c)=-\frac{c}{3}-2a,$$

we have got

$$\frac{8c}{3}+16a-12a+16a-4c-24a-4c=0,$$

so that

$$\frac{8c}{3}-4a=0,$$

which has a solution

$$c=\frac{3a}{2}.$$

Thus,

$$c=\frac{3\ P}{(2)\ 4!\ EI},$$

so that

$$M=-2cEI=-\frac{3\ P}{4!}.$$

4.2. Checking this last result for M by solving the concerning ordinary differential equation

In this section we check the result obtained for M by solving the following ODE,

$$EIw_{,xxxx}-P=0,\ \text{in}\ \Omega=[0,1],$$

with the boundary conditions,

$$w(0)=w(1)=w_{,x}(0)=w_{,xx}(1)=0.$$

We recall the general solution stands for

$$w(x)=ax^4+bx^3+cx^2+dx+e,$$

where

$$a=\frac{P}{4!EI}.$$

From $w(0)=0$, we obtain $e=0$.
From $w_{,x}(0)=0$, we obtain $d=0$.
From $w(1)=0$, we have

$$a+b+c=0.$$

From $w_{xx}(1)=0$, we have

$$12a+6b+2c=0.$$

From such results, we obtain

$$b=-\frac{5a}{2},$$

and
$$c = \frac{3a}{2},$$

Observe that
$$M(x) = -EIw_{,xx},$$
so that in particular
$$M = M(0) = -EIw_{,xx}(0) = -EI(2c) = -\frac{3\ P}{4!}.$$

This value for M here obtained coincide with the one obtained in the previous subsection, as expected. The objective of this section is complete.

Appendix A

In this appendix we present the statement of the Implicit Function Theorem in Banach Spaces.

Theorem 5 (The implicit function theorem)**.** *Let V, U, W be Banach spaces. Let $F : V \times U \to W$ be a functions such that*
$$F(x_0, u_0) = \mathbf{0},$$
where $(x_0, u_0) \in V \times U$.

Assume there exists $r > 0$ such that F is Fréchet differentiable and $F_x(x, u)$ is continuous in (x, u) in $B_r(x_0, u_0)$. Suppose also $[F_x(x_0, u_0)]^{-1}$ exists and is bounded so that there exists $\rho > 0$ such that
$$0 < \|[F_x(x_0, u_0)]^{-1}\| \leq \rho.$$

Under such hypotheses, there exist $0 < \varepsilon_1 < r/2$ and $0 < \varepsilon_2 < 1$ such that for each $u \in B_{\varepsilon_1}(u_0)$, there exists $x \in B_{\varepsilon_2}(x_0)$ such that
$$F(x, u) = \mathbf{0},$$
where we denote $x = x(u)$ so that
$$F(x(u), u) = \mathbf{0}.$$

Moreover, there exists $\delta > 0$ such that $0 < \delta\rho < 1$, such that for each $u, v \in B_{\varepsilon_1}(u_0)$ we have
$$\|x(u) - x(v)\| \leq \frac{\rho^2\delta}{1 - \rho\delta}\|F(x(v), u) - F(x(v), v)\|.$$

Finally, if there exists $K > 0$ such that
$$\|F_u(x, u)\| \leq K,\ \forall (x, u) \in B_{\varepsilon_2}(x_0) \times B_{\varepsilon_1}(u_0),$$
so that
$$\|F(x, u) - F(x, v)\| \leq K\|u - v\|,\ \forall (x, u) \in B_{\varepsilon_2}(x_0) \times B_{\varepsilon_1}(u_0),$$
then
$$\|x(u) - x(v)\| \leq K_1\|u - v\|,$$
where
$$K_1 = K\frac{\rho^2\delta}{1 - \delta\rho}.$$

We present also a respective corollary.

Corollary 1. *Consider the hypotheses and statements of the last theorem. Moreover, assume $F_x : V \times U \to W$ is such that $[F_x(x, u)]^{-1}$ exists and it is bounded in $B_r(x_0, u_0)$.*

Suppose also, F is Fréchet differentiable in $B_r(x_0, u_0)$.

Let $\varphi \in U$.

Under such hypotheses,

$$x'(u,\varphi) = -[F_x(x(u),u)]^{-1}[F_u(x(u),u)](\varphi),$$

where

$$x'(u,\varphi) = \lim_{t\to 0}\frac{x(u+t\varphi)-x(u)}{t}.$$

From this last corollary we may also obtain the Fréchet differentiability of $x(u)$ in a concerning neighborhood of u_0.

5. Conclusion

In this article, we have presented a formal proof Castigliano Theorem in a linear elasticity theory context.

We have also presented a generalization of such a result for a non-linear elasticity context and a numerical example to illustrate its applicability.

Acknowledgments: The author is very grateful to an anonymous referee whose the comments and suggestions helped me improve a lot the quality of this article.

References

[1] Botelho, F. (2025). A Formal Proof of Castigliano Theorem and a Related Generalization.

[2] Beer, F. P., Johnston, E. R., DeWolf, J. T., & Mazurek, D. F. (2015). *Mechanics of Materials*. 7th_edition. New York. MeGraw-Hill Education Ltd, 1.

[3] Hartmann, F. (1982). Castigliano'Theorem and its Limits. *ZAMM-Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik, 62*(12), 645-650.

[4] Steiner, W. (2014). The use of Castigliano's theorem in Coulomb friction problems. *Acta Mechanica, 225*(9), 2471-2483.

[5] Marotti de Sciarra, F. (1997). Direct and dual theorems of Castigliano and Clapeyron for generalized elastic models. *Acta Mechanica, 124*(1), 107-130.

[6] Yoo, C. H., & Lee, S. (2011). *Stability of Structures: Principles and Applications*. Elsevier.

[7] Pippard, A. J. S., & Baker, J. F. (1943). *The Analysis of Engineering Structures*. London.

[8] Adams, R. A., & Fournier, J. J. (2003). *Sobolev Spaces* (Vol. 140). Elsevier.

[9] Robinson, S. M. (1980). Strongly regular generalized equations. *Mathematics of Operations Research, 5*(1), 43-62.

[10] Ito, K., & Kunisch, K. (2008). *Lagrange Multiplier Approach to Variational Problems and Applications*. Society for Industrial and Applied Mathematics.

[11] Botelho, F. S. (2021). *Functional Analysis, Calculus of Variations and Numerical Methods for Models in Physics and Engineering*. CRC Press, Taylor & Francis Group.