London '13

Tu-07-01

Nonlinear partial Functional Derivative and Nonlinear LS Seismic Inversion

R.S. Wu* (University of California), L. Ye (University of California, Santa Cruz) & Y. Zheng (Massachussets Inst. of Technology)

SUMMARY

Taylor expansion of the full nonlinear partial derivative (NLPD) operator is directly related to the full scattering series (Born series) which has a serious convergence problem for strong scattering. The renormalization procedure applied to the Taylor-Fréchet series leads to the De wolf approximation of NLPD, which changes the Fredholm type series into a Volterra type series so that renormalized Fréchet series has a guaranteed convergence. Numerical simulations demonstrate the different convergence behaviors of the two types of series for NLPD in the case of strong perturbations. Preliminary study on the LS inversion theory using the nonlinear kernel leads to an inversion scheme of simultaneous updating both the model parameters and propagators.

Introduction

The Fréchet derivative is widely used in geophysical inverse problems. We know that the linear Fréchet derivative corresponds to the Born modeling (Tarantola, 1984; Pratt et al., 1998), For the real Earth, the wave equation is strongly nonlinear with respect to the medium parameter changes. Therefore, it is interesting to see how the higher order terms of the nonlinear functional derivatives, i.e. the higher order Fréchet derivatives influence the inversion procedure and its convergence. Wu and Zheng (2012) introduced the higher order Fréchet derivatives for the acoustic wave equation, and termed the full functional derivative as the "nonlinear Fréchet derivative". Because of the possible confusion with the conventional notation, in which the "Fréchet derivative" is identical with "linear Fréchet derivative", here we name the sum of all terms as "*nonlinear functional derivative*" or "*nonlinear partial derivative* (NLPD)". In this paper we will discuss the renormalization of the Taylor-Fréchet series into a De Wolf-Fréchet series and then compare the convergence properties of these different Fréchet series with numerical demonstration. Finally we point out point out the potential of application of nonlinear partial derivative to least-square waveform inversion.

Higher order Fréchet derivatives and nonlinear partial derivative operator

We write the forward problem into an operator form,

$$\mathbf{d} = \mathbf{A}(\mathbf{m}) \tag{1}$$

where **d** is the data vector (pressure field in the case of acoustic wave equation), **m** is the model vector, and **A** is the forward modeling operator. Assume an initial model $\mathbf{m_0}$**,** we want to quantify the sensitivity of the data change δ**d** (also called "data residual") to the model perturbation δ**m,**

$$\delta\mathbf{d} = \mathbf{d} - \mathbf{d}_0 = \mathbf{F}_{\mathbf{m}_0}(\delta\mathbf{m}) = \mathbf{A}(\mathbf{m}_0 + \delta\mathbf{m}) - \mathbf{A}(\mathbf{m}_0) \tag{2}$$

We know that $\mathbf{F}(\delta\mathbf{m})$ is a nonlinear differential operator, and can be expanded into Taylor expansion at the current model (Kwon and Yazici, 2010; Wu and Zheng, 2012):

$$\mathbf{F}(\delta\mathbf{m}) = A'(\mathbf{m}_0)\delta\mathbf{m} + \frac{1}{2!}A''(\mathbf{m}_0)(\delta\mathbf{m})^2 + \cdots + \frac{1}{n!}A^{(n)}(\mathbf{m}_0)(\delta\mathbf{m})^n + \cdots \tag{3}$$

where $\mathbf{A}'$, $\mathbf{A}''$, and $\mathbf{A}^{(n)}$ are the first, second, and the nth order Fréchet derivatives. It has been shown that in the wave equation case, the higher order Fréchet derivatives can be realized by consecutive applications of the scattering operator and a zero-order propagator to the source, and the Taylor-Fréchet derivative series is closely related to the Born scattering series (Wu and Zheng, 2012):

$$\mathbf{A}^{(n)}(\mathbf{m}_0)(\delta\mathbf{m})^n = n![G_0\mathbf{S}\varepsilon]^n p_0 = n!G_0(\mathbf{S}_n\varepsilon G_0\mathbf{S}_{n-1}\varepsilon......G_0\mathbf{S}_1\varepsilon)p_0 \tag{4}$$

where p_0 is the incident field, g_0 is the background Green's function, $\mathbf{G}_0$ is the background Green's operator and $\mathbf{S}$ is a local scattering operator (scattering pattern). If we define a *nonlinear partial derivative operator* $\mathbf{A}^{(NLPD)}(\mathbf{m}_0,\delta\mathbf{m})$ based on the nonlinear differential operator $\mathbf{F}(\mathbf{m}_0,\delta\mathbf{m})$ through $\mathbf{A}^{(NLPD)}(\mathbf{m}_0,\delta\mathbf{m})\delta\mathbf{m} = \mathbf{F}(\mathbf{m}_0,\delta\mathbf{m})$, then we have

$$\mathbf{A}^{(NLPD)}(\mathbf{m}_0,\delta\mathbf{m}) = A'(\mathbf{m}_0) + \frac{1}{2!}A''(\mathbf{m}_0)(\delta\mathbf{m}) + \cdots + \frac{1}{n!}A^{(n)}(\mathbf{m}_0)(\delta\mathbf{m})^{n-1} + \cdots. \tag{5}$$

Note that $\mathbf{A}^{(NLPD)}$ is $\delta\mathbf{m}(\mathbf{x})$-dependent because of the nonlinear mutual interactions (multiple scattering) between perturbations. Compare with the traditional *linear perturbation model*

$$\mathbf{F}(\delta\mathbf{m}) = A'(\mathbf{m}_0)\delta\mathbf{m}, \tag{6}$$

we call equation (5) with the operator (or kernel) $\mathbf{A}^{(NLPD)}(\mathbf{m}_0,\delta\mathbf{m})$ as *nonlinear perturbation model*.

Renormalization of the Taylor-Fréchet derivatives series and the De wolf approximation

If we split the scattering operator into forward scattering and backscattering parts

$$\mathbf{S} = \mathbf{S}^f + \mathbf{S}^b \tag{7}$$

and substitute it into the Fréchet series, we can have all combinations of higher order forward and backward derivatives. The De Wolf approximation in scattering series corresponds to neglecting multiple backscattering (reverberations), i.e. dropping all the terms containing two or more backscattering operators but keeping all the forward scattering terms untouched (De Wolf, 1971, 1985; Wu, 1994, 2003; Wu et al., 2012).

To demonstrate the principle of the nonlinear partial derivative, we treat a simple problem of transmission tomography in smooth media. In this case, there is no reflection and we have only forward scattering due to velocity perturbations. The higher order Fréchet derivatives only involve forescattering operator $\mathbf{S}^f$, and its application to the model perturbation yields,

$$\begin{aligned}\frac{1}{n!}\mathbf{A}_f^{(n)}(\mathbf{m}_0)\delta\mathbf{m}^n &= G_0\left(\mathbf{S}_n^f\delta\mathbf{m}G_0\mathbf{S}_{n-1}^f\delta\mathbf{m}...G_0\mathbf{S}_{i+1}^f\delta\mathbf{m}G_0\mathbf{S}_i^f\delta\mathbf{m}(\mathbf{x})G_0\mathbf{S}_{i-1}^f\delta\mathbf{m}....G_0\mathbf{S}_1^f\delta\mathbf{m}\right)g_0 \\ &= G_0(\mathbf{S}^f\delta\mathbf{m}G_0)^{n-i}\mathbf{S}_i^f\delta\mathbf{m}(\mathbf{x})(G_0\mathbf{S}^f\delta\mathbf{m})^{i-1}g_0\end{aligned} \tag{8}$$

Following the renormalization procedure in the De wolf approximation (De Wolf, 1985; , Delamotte**,** 2004; Wu, 2003; Wu et al., 2012), we sum up all the higher order terms in the Taylor series firstly for the multiple forescattering operators on the left-hand side of $\delta\mathbf{m}(\mathbf{x})$ (receiver path) and then for that on the right-hand side of $\delta\mathbf{m}(\mathbf{x})$ (source path) in the above equation, resulting in

$$G_f^{n-i} = \sum_{l=0}^{n-i} G_0\left(\mathbf{S}^f\delta\mathbf{m}G_0\right)^l, \qquad g_f^i = \sum_{l=0}^{i-1}\left(G_0\mathbf{S}^f\delta\mathbf{m}\right)^l g_0 \tag{9}$$

When $n\to\infty$ and the step length becomes infinitely small, we reach the renormalized G_f and g_f. Under this approximation, the Taylor series of NLPD is renormalized to

$$A_f^{NLPD}\left(\mathbf{m}_0,\delta\mathbf{m}\right)\delta\mathbf{m}(\mathbf{x}) = G_f\mathbf{S}^f\delta\mathbf{m}(\mathbf{x})g_f \tag{10}$$

We see that when NLPD applying to a perturbation function under the De Wolf approximation, all the nonlinear interactions due to forescattering are incorporated into g_f and G_f.

Convergence property of the renormalized Fréchet series (NLPD) under the De Wolf approximation

Compared with the Taylor expansion (Born series) the De Wolf-Fréchet series has the stability and efficiency advantages. First look at the stability (series convergence) problem. The original Taylor series (Born series), derived by applying the Born-Neumann iterative procedure to the Lippmann-Schwinger equation, is a Fredholm type, and has the well-known problem of limited region of convergence (slow convergence and divergence). The iterative procedure based on the series using gradient or Newton method will have no guarantee of converging to a correct solution. In contrast, if we write out explicitly the nth term of g_f in (10) (same for G_f) in the infinite series (9)

$$g_f = \sum_{l=0}^{\infty}\left(G_0\mathbf{S}^f\delta\mathbf{m}\right)^l g_0 = g_0 + g_1 + g_2 + ...g_n + ...$$

$$\begin{cases} g_0 = g_0(\mathbf{x},\mathbf{x}_s), \quad ...... \\ g_n = \left(G_0\mathbf{S}^f\delta\mathbf{m}\right)^n g_0 = \mathbf{G}_0(\mathbf{x}_r,\mathbf{x}_n)\mathbf{S}^f\delta\mathbf{m}(\mathbf{x}_n)\mathbf{G}_0\mathbf{S}^f\delta\mathbf{m}(\mathbf{x}_{n-1})\cdots\cdots\mathbf{G}_0\mathbf{S}^f\delta\mathbf{m}(\mathbf{x}_1)g_0(\mathbf{x}_1,\mathbf{x}_s) \\ \quad = \int_V dz_n dy_n dx_n g_0(\mathbf{x};\mathbf{x}_n)\mathbf{S}^f\delta\mathbf{m}(\mathbf{x}_n)... \\ \quad \times\int_V dz_2 dy_2 dx_2 g_0(\mathbf{x}_3;\mathbf{x}_2)\mathbf{S}^f\delta\mathbf{m}(\mathbf{x}_2)\int_V dz_1 dy_1 dx_1 g_0(\mathbf{x}_2;\mathbf{x}_1)\mathbf{S}^f\delta\mathbf{m}(\mathbf{x}_1)g_0(\mathbf{x}_1,\mathbf{x}_s) \\ \qquad\qquad , \quad x > x_n > x_{n-1} > ...x_2 > x_1 > x_s \end{cases} \tag{11}$$

where x is taken as the forward marching direction. We see that each term of the series in (11) is a Volterra type integral (Tricomi, 1985; Schetzen, 1980). Therefore, series in (11) is a Volterra series which converges absolutely and uniformly (ibid). Therefore, the series for $A_f^{NLPD}\left(\mathbf{m}_0,\delta\mathbf{m}\right)$ has a guaranteed convergence. Physically, the renormalization procedure can be understood as a way of rearranging the order of mutual cancelations (destructive interference) between different terms. In Born series, each term involves multiple whole volume integral and the mutual cancellations only act between each terms in the final stage. So forward scattering for the whole volume may become very strong or singular for each term. If each individual term does not blow up in the process, then the final summation may get an approximate solution due to mutual cancellations. However, if the errors of the individual terms become too big, the final summation may blow up or give a wrong result. In comparison, the mutual cancellations for the renormalized series (11) are realized step-by-step during the forward marching process so that the forward-scattering accumulation is not allowed to develop into a full-blown catastrophe.

Now we show some examples to compare the convergence of the Taylor-Fréchet series and the renormalized Fréchet series. The source is located on the top and receivers are distributed along the bottom as shown on the upper-right panel in Figure 1. Source has a Ricker wavelet, centered at 20 Hz ($\lambda = 100m$). The model is a fast Gaussian anomaly $dv(x,z) = \varepsilon v_0 \exp\left[-r^2/\left(2a^2\right)\right]$, embedded in a constant background $v_0 = 2$ km/s . The Gaussian ball has parameters $a = 3\lambda$, and the perturbation is given as $\varepsilon = 5\%,\ 8\%\ \&\ 10\%$, respectively. We see that for weak scattering (5%) the Born series converges very fast, in this case only 10 terms; For medium-strength perturbations, it converges slowly (17 terms); However, for strong scattering ($a = 3\lambda$, $\varepsilon = 10\%$), the series diverges! This demonstrates the limited region of convergence for the Born series and the related iteration process. In comparison, we plot the corresponding results using the De wolf approximation in Figure 2. We see that it has a guaranteed convergence and the results are nearly the same as the FD simulation results. The other advantage of NLPD in the form of (10) is the efficiency. Although it is in the form of an infinite series, it can be implemented efficiently by the thin-slab propagator or GSP (generalized screen propagator) (Wu, 1994, 2003), which is a one sweep algorithm for forward scattering problem. The equivalence of the multiple forward scattering series and the thin-slab propagator has been proved (Wu et al., 2012).

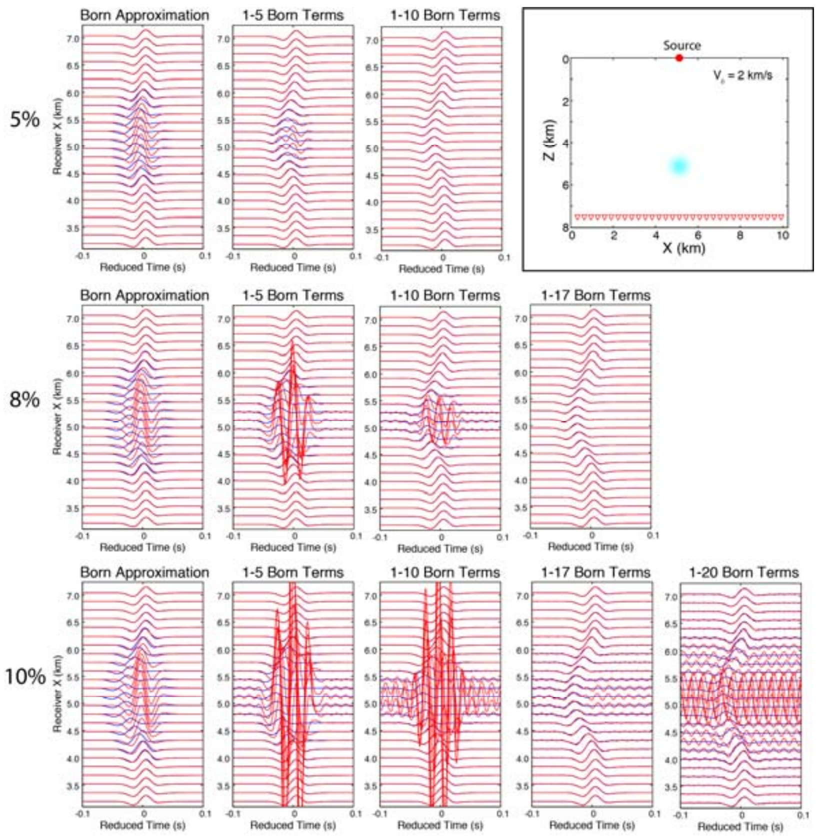

Figure 1 Taylor-Fréchet (Born) series test for a Gaussian ball with a=3λ, $\varepsilon = 5\%,\ 8\%\ \&\ 10\%$. Blue wiggles are from FD calculations; red: Born series summing up to certain orders.

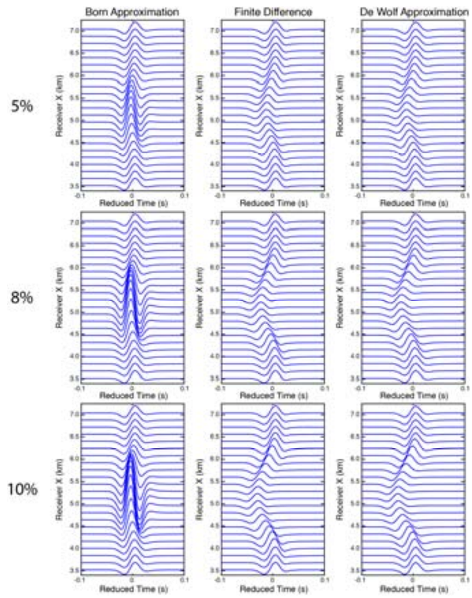

Figure 2 Comparison of results from Born approximation (linear Fréchet derivative) (left), FD (mid) and De Wolf (right) for $\varepsilon = 5\%$, 8% & 10%.

Conclusion and discussion

Taylor expansion of the full nonlinear partial derivative (NLPD) operator is directly related to the full scattering series (Born series) which has a serious convergence problem for strong scattering. The renormalization procedure applied to the Taylor-Fréchet series leads to the De wolf approximation of NLPD, which has a guaranteed convergence. Preliminary study on the LS inversion theory using the NLPD kernel leads to an inversion scheme of simultaneous updating both the model and propagator which may reduce the initial model dependence of seismic inversion.

Acknowledgement

The research is supported by the WTOPI (Wavelet Transform on Propagation and Imaging for seismic exploration) Research Consortium at the University of California, Santa Cruz.

References

Delamotte, **.,** 2004. A hint of renormalization, *Am. J. of Phys.*, **72, No 2,** 170-184.

De Wolf, D.A., 1985. Renormalization of EM fields in application to large-angle scattering from randomly continuous media and sparse particle distributions. *IEEE Trans. Antennas and Propagations* **AP-33**, 608–615.

Kwon, K. and Yazici, B., 2010. Born expansion and Frechet derivatives in nonlinear diffuse optical tomography, *Computers and Mathematics with Applications*, doi: 10.1016/j.camwa.2009.07.088.

Schetzen, M., 1980, *The Volterra and Wiener Theories of Nonlinear Systems*, Wiley.

Tarantola, A., 1984. Inversion of seismic-reflection data in the acoustic approximation, *Geophysics*, **49**(8), 1259-1266.

Tricomi, F.G., 1985, Integral Equations, Dover Publications, Inc..

Wu, R.S., 2003. Wave propagation, scattering and imaging using dual-domain one-way and one-return propagators, *Pure and Appl. Geophys.*, **160**(3/4), 509-539.

Wu, R.S**.,** Xie, X.B. and Jin, S., 2012. One-Return Propagators and The Applications in Modeling and Imaging, Chapter 2 in "*Imaging, Modeling and Assimilation in Seismology",* Higher Education Press Limited Company, Beijing, 65-105.

Wu, R.S. and Y. Zheng, 2012. Nonlinear Fréchet derivative and its De Wolf approximation, Expanded Abstracts of Society of Exploration Gephysicists, SI 8.1.