International Scholarly Research Network
ISRN Applied Mathematics
Volume 2012, Article ID 539359, 10 pages
doi:10.5402/2012/539359

Research Article

Green's Theorem for Sign Data

Louis M. Houston

The University of Louisiana at Lafayette, Lafayette, LA 70504-4210, USA

Correspondence should be addressed to Louis M. Houston, houston@louisiana.edu

Received 14 March 2012; Accepted 19 April 2012

Academic Editors: A. Bellouquid and M.-H. Hsu

Sign data are the signs of signal added to noise. It is well known that a constant signal can be recovered from sign data. In this paper, we show that an integral over variant signal can be recovered from an integral over sign data based on the variant signal. We refer to this as a generalized sign data average. We use this result to derive a Green's theorem for sign data. Green's theorem is important to various seismic processing methods, including seismic migration. Results in this paper generalize reported results for 2.5D data volumes in which Green's theorem applies to sign data based only on traditional sign data recovery.

1. Introduction

In certain cases, data, consisting of coherent signal and random noise, is summed repeatedly in order to enhance the signal and reduce the noise. It has been found that when the signal-to-noise ratio is between 0.1 and 1, if only the signs of the data are retained prior to summation, then the signal can be recovered [1]. We refer to this type of data as sign data. Sign data has advantages in significantly reducing the amount of space needed to record the data. The reduction in space can amount to a ratio of 20 to 1 bits of data storage [1].

In the seismic industry, it has been discovered that most of the data processing methods used on regular data are also effective on sign data [2]. Many of these data processing methods are dependent on Green's theorem. In particular, Kirchhoff migration is dependent on Green's theorem. Migration is a seismic process that corrects data coordinates which are distorted by attributes of the seismic experiment. In accordance with Green's theorem, we show that an integral of an operation on sign data over a volume is equal to an integral of an operation of the original data over a surface.

This result generalizes the work of Houston and Richard [3] in which, based only on traditional sign data recovery, it was shown that Green's theorem is satisfied when the sign data encompasses a 2.5D volume. 2.5D data is data which is three-dimensional but only has

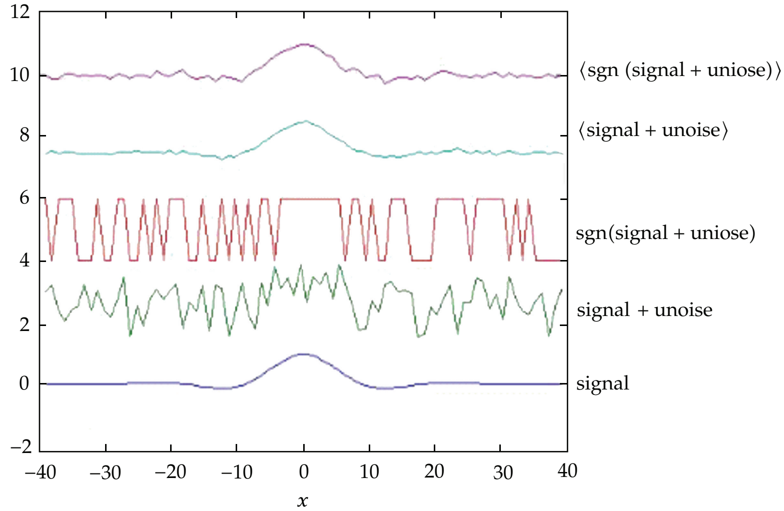

Figure 1: A synthetic signal $(\sin c(x/3)e^{-(x/15)^2})$ and the computer-generated uniform random noise used to examine sign-bit amplitude recovery. This test has noise of unit magnitude ($a = 1$) and a signal-to-noise ratio of one. Shown from bottom to top is the signal, the signal plus noise, the sign of the signal plus noise, the average over 200 iterations of signal plus noise, and the average over 200 iterations of the sign of signal plus noise.

variations along two dimensions. In this paper, we find that Green's theorem is satisfied for sign data when the data volume is largely symmetric. For cases in which the data volume is of arbitrary shape we have derived a variant of Green's theorem.

2. A Generalized Sign Data Average

It was shown by O'Brien et al. [1] that there is a recovery of amplitude from sign data in the presence of uniform random noise. If the random noise X_j has amplitude a, the signal is designated by the function f_k, and the number of iterations is $M = \sum_j$, then we can write

$$f_k = \frac{a}{M}\sum_j \operatorname{sgn}(f_k + X_j), \tag{2.1}$$

where

$$\operatorname{sgn}(y) = \begin{cases} +1, & y > 0 \\ 0, & y = 0 \\ -1, & y < 0. \end{cases} \tag{2.2}$$

Figure 1 illustrates the data recovery process including sign data.

Multiply both sides of (2.1) by the function g_k:

$$g_k f_k = \frac{a}{M} \sum_j g_k \operatorname{sgn}\left(f_k + X_j\right). \tag{2.3}$$

Now sum both sides of (2.3) over the k index:

$$\sum_k g_k f_k = \frac{a}{M} \sum_k \sum_j g_k \operatorname{sgn}\left(f_k + X_j\right). \tag{2.4}$$

If we allow $k \rightarrow j$, then (2.4) becomes

$$\sum_j g_j f_j = \frac{a}{M} \sum_j \sum_j g_j \operatorname{sgn}\left(f_j + X_j\right) \tag{2.5}$$

or

$$\sum_j g_j f_j = a \sum_j g_j \operatorname{sgn}\left(f_j + X_j\right). \tag{2.6}$$

It is clear that if $f_j \rightarrow f$ and $g_j \rightarrow g$, then (2.6) becomes

$$f = \frac{a}{M} \sum_j \operatorname{sgn}\left(f + X_j\right), \tag{2.7}$$

which is essentially (2.1). A continuous version of (2.6) might be written as

$$\int g(v) f(v) dv = a \int g(v) \operatorname{sgn}(f(v) + v) dv. \tag{2.8}$$

An argument for the consistency of (2.8) is as follows. Let $f(v) \rightarrow f$. Then we have

$$f = \frac{a \int g(v) \operatorname{sgn}(f + v) dv}{\int g(v) dv}. \tag{2.9}$$

Now integrate over all values of v:

$$f = \frac{a \int_{-\infty}^{\infty} g(v) \operatorname{sgn}(f + v) dv}{\int_{-\infty}^{\infty} g(v) dv}. \tag{2.10}$$

Let $g(v)$ be a uniform probability density, $\rho(v)$.

That implies

$$\int_{-\infty}^{\infty} g(v)dv = 1, \tag{2.11}$$

and (2.10) becomes

$$f = a\int_{-\infty}^{\infty} \rho(v)\,\mathrm{sgn}(f+v)dv, \tag{2.12}$$

which was shown by Houston et al. [4].

3. A Generalized Average for Sign Data Derivatives

Consider the following nth order forward finite difference [5]:

$$\Delta_h^n f(v) = \sum_{i=0}^{n}(-1)^i \binom{n}{i} f(v+(n-i)h), \tag{3.1}$$

with $\binom{n}{i} = n!/i!(n-i)!$.

If we make the variable, v discrete by choosing a small real interval, q and writing

$$f_j = f(jq), \tag{3.2}$$

where j is an integer index, then (3.1) becomes

$$\Delta_h^n f_j = \sum_{i=0}^{n}(-1)^i \binom{n}{i} f(jq+(n-i)h). \tag{3.3}$$

Because the finite difference is a linear operator, we can use (2.6) to derive

$$\sum_j g_j \Delta_h^n f_j = a\sum_j g_j \Delta_h^n\,\mathrm{sgn}(f_j + X_j). \tag{3.4}$$

Equation (3.4) takes into account the fact that

$$\sum_j g(jq)f(jq+lh) = a\sum_j g(jq)\,\mathrm{sgn}(f(jq+lh)+X_j), \tag{3.5}$$

where l is an arbitrary index.

Since

$$\frac{d^n}{dv^n} = \lim_{h\to 0}\frac{\Delta_h^n}{h^n}, \tag{3.6}$$

(3.4) suggests the following equation:

$$\int g(v)\frac{d^n}{dv^n}f(v)dv = a\int g(v)\frac{d^n}{dv^n}\operatorname{sgn}\bigl(f(v)+v\bigr)dv. \tag{3.7}$$

Let $d^n/dv^n \to d^n/du^n$ and $f(v) \to f(u)$.
Then from (3.7) we have

$$\frac{d^n}{du^n}f(u) = \frac{a\int g(v)(d^n/du^n)\operatorname{sgn}\bigl(f(u)+v\bigr)dv}{\int g(v)dv}. \tag{3.8}$$

Once again, integrate over all v, let $g(v) = \rho(v)$, and (3.8) becomes

$$\frac{d^n}{du^n}f(u) = a\int_{-\infty}^{\infty}\rho(v)\operatorname{sgn}\bigl(f(u)+v\bigr)dv, \tag{3.9}$$

which was shown by Houston et al. [4].

4. The Application to Green's Theorem

Equation (3.7) implies the special case:

$$\int g(v)\frac{d^2}{dv^2}f(v)dv = a\int g(v)\frac{d^2}{dv^2}sgn\bigl(f(v)+v\bigr)dv. \tag{4.1}$$

Employing three variables in (4.1) yields

$$\begin{aligned}&\int g(v_1,v_2,v_3)\frac{\partial^2}{\partial v_i^2}f(v_1,v_2,v_3)dv_1dv_2dv_3\\&\quad= a\int g(v_1,v_2,v_3)\frac{\partial^2}{\partial v_i^2}\operatorname{sgn}\bigl(f(v_1,v_2,v_3)+v_i\bigr)dv_1dv_2dv_3,\end{aligned} \tag{4.2}$$

which can be simplified if $g = g(v_1,v_2,v_3)$, $f = f(v_1,v_2,v_3)$, and $dV = dv_1dv_2dv_3$:

$$\int g\frac{\partial^2}{\partial v_i^2}f\,dV = a\int g\frac{\partial^2}{\partial v_i^2}\operatorname{sgn}(f+v_i)dV. \tag{4.3}$$

Summation over the variables yields

$$\sum_i \int g \frac{\partial^2}{\partial v_i^2} f \, dV = a \sum_i \int g \frac{\partial^2}{\partial v_i^2} \operatorname{sgn}(f + v_i) dV. \tag{4.4}$$

Consequently, (4.4) can be written as

$$\int g \nabla^2 f \, dV = a \int g \sum_i \frac{\partial^2}{\partial v_i^2} \operatorname{sgn}(f + v_i) dV. \tag{4.5}$$

If we make the mapping

$$g(v) \longrightarrow \frac{\partial^2}{\partial v_i^2} g, \tag{4.6}$$

then (2.8) becomes

$$\int f \frac{\partial^2}{\partial v_i^2} g \, dV = a \int \operatorname{sgn}(f + v_i) \frac{\partial^2}{\partial v_i^2} g \, dV. \tag{4.7}$$

Summation over the variables yields

$$\sum_i \int f \frac{\partial^2}{\partial v_i^2} g \, dV = a \sum_i \int \operatorname{sgn}(f + v_i) \frac{\partial^2}{\partial v_i^2} g \, dV, \tag{4.8}$$

which can be written as

$$\int f \nabla^2 g \, dV = a \int \sum_i \operatorname{sgn}(f + v_i) \frac{\partial^2}{\partial v_i^2} g \, dV. \tag{4.9}$$

Subtracting (4.9) from (4.5) yields

$$\int \left(g \nabla^2 f - f \nabla^2 g \right) dV = a \int \left(g \sum_i \frac{\partial^2}{\partial v_i^2} \operatorname{sgn}(f + v_i) - \sum_i \operatorname{sgn}(f + v_i) \frac{\partial^2}{\partial v_i^2} g \, dV \right). \tag{4.10}$$

Green's theorem is

$$\oint_S \left(U_1 \frac{\partial U_2}{\partial n} - U_2 \frac{\partial U_1}{\partial n} \right) dS = \int_V \left(U_1 \nabla^2 U_2 - U_2 \nabla^2 U_1 \right) dV. \tag{4.11}$$

Consequently, we can write a variant of Green's theorem for sign data as

$$\oint_S \left(g \frac{\partial f}{\partial n} - f \frac{\partial g}{\partial n} \right) dS = a \int \left(g \sum_i \frac{\partial^2}{\partial v_i^2} \operatorname{sgn}(f + v_i) - \sum_i \operatorname{sgn}(f + v_i) \frac{\partial^2}{\partial v_i^2} g \right) dV. \tag{4.12}$$

If the spatial intervals are uniform, that is, $v_1 \in [a_1, b_1]$, $v_2 \in [a_2, b_2]$, $v_3 \in [a_3, b_3]$ and $[a_1, b_1] = [a_2, b_2] = [a_3, b_3]$, then (4.12) becomes

$$\oint_S \left(g\frac{\partial f}{\partial n} - f\frac{\partial g}{\partial n} \right) dS = a \int \left(g\nabla^2 \operatorname{sgn}(f + v_i) - \operatorname{sgn}(f + v_i)\nabla^2 g \right) dV. \tag{4.13}$$

In the case of (4.13), we see that Green's theorem is satisfied by replacing the function f with sign data and dividing the surface integral by the noise amplitude. Therefore, all processing of real data based on Green's theorem will be effective for sign data with an associated variance. We also note that the effectiveness of Green's theorem on sign data is enhanced by operating over a symmetric volume. When the volume is not symmetric, we can apply the variant of Green's theorem given by (4.12).

5. The 2.5D Case

Now let us consider the special case for which the functions f and g have only two-dimensional variation. That is,

$$f \longrightarrow f(v_1, v_2) g \longrightarrow g(v_1, v_2). \tag{5.1}$$

Without using a generalized sign data average, this implies

$$a \int \left(g\nabla^2 \operatorname{sgn}(f + v_i) - \operatorname{sgn}(f + v_i)\nabla^2 g \right) dV = (b_3 - a_3) \int \left(g\nabla^2 f - f\nabla^2 g \right) dv_1 dv_2 \tag{5.2}$$

or

$$\oint_S \left(g\frac{\partial f}{\partial n} - f\frac{\partial g}{\partial n} \right) dS = (b_3 - a_3) \int \left(g\nabla^2 f - f\nabla^2 g \right) dv_1 dv_2. \tag{5.3}$$

Equation (5.3) is Green's theorem for sign data when the data encompasses a 2.5D volume and is consistent with results reported in Houston and Richard [3].

6. Computational Tests

Equation (2.6) derives from a weighted average of the function $y_j = \operatorname{sgn}(f_j + X_j)$. We can thus write the expectation value of y_i as

$$E(Y) = \frac{\sum_j g_j y_j}{\sum_j g_j}. \tag{6.1}$$

We find that

$$E(Y) = \frac{\sum_j g_j f_j}{a \sum_j g_j}. \tag{6.2}$$

The variance, $\mathrm{Var}(Y)$, can be written as

$$\mathrm{Var}(Y) = E(Y - E(Y))^2. \tag{6.3}$$

This can be reduced to a simpler form by using the fact that

$$E\left(Y^2\right) = 1. \tag{6.4}$$

Thus, we have

$$\mathrm{Var}(Y) = 1 - (E(Y))^2. \tag{6.5}$$

We can demonstrate (2.6),

$$\sum_j g_j f_j = a \sum_j g_j \operatorname{sgn}\left(f_j + X_j\right), \tag{6.6}$$

computationally. Let

$$\begin{aligned} f_j &= \sin(jq), \\ g_j &= \cos(jq). \end{aligned} \tag{6.7}$$

Consequently, we want to demonstrate that

$$\sum_j \cos(jq) \sin(jq) = a \sum_j \cos(jq) \operatorname{sgn}\left(\sin(jq) + X_j\right). \tag{6.8}$$

We can compute the average percentage error β as

$$\beta = 100 \sqrt{\frac{\left\langle \left(\sum_j \cos(jq) \sin(jq) - a \sum_j \cos(jq) \operatorname{sgn}\left(\sin(jq) + X_j\right)\right)^2 \right\rangle}{\left(\sum_j \cos(jq) \sin(jq)\right)^2}}. \tag{6.9}$$

Table 1: Twenty trials which compare the values of the indicated expressions for M iterations per trial. The trials are divided into ten trials for two different noise amplitudes. Recall that $M = \sum_j$. Let $M = 10000$ and $q = 0.0001$.

$\sum_{j=1}^{M} \cos(jq) \sin(jq)$	$a \sum_{j=1}^{M} \cos(jq)\,\mathrm{sgn}(\sin(jq) + X_j)$	
$a = 1$		
3540.594404	3529.425407	(1)
↓	3569.469208	(2)
	3460.5978	(3)
	3540.394914	(4)
	3542.955134	(5)
	3644.993387	(6)
	3441.978126	(7)
	3510.495952	(8)
	3544.098984	(9)
	3645.64803	(10)
$\beta = 0.56447\%, \mathrm{Var}(Y) = 0.822949$		
$a = 2$		
3540.594404	3578.540737	(1)
↓	3468.32837	(2)
	3063.26568	(3)
	3629.793227	(4)
	3562.692845	(5)
	3605.376332	(6)
	3284.005063	(7)
	3375.67384	(8)
	3356.467161	(9)
	3500.608956	(10)
$\beta = 1.731154\%, \mathrm{Var}(Y) = 0.955737$		

We should see a correlation between β and $\mathrm{Var}(Y)$. The results of this comparison are shown in Table 1.

7. Conclusions

Using the results of sign data signal recovery leads to a derivation of a generalized sign data average. Extending these results to incorporate derivatives leads to a variant of Green's theorem for sign data. We find that Green's theorem directly applies to sign data when the data volume is symmetric and the surface integral is divided by the noise amplitude. A specific application of this result is that Green's theorem applies to sign data when the data volume is 2.5D and a generalized sign data average is not required.

Acknowledgment

Discussions with Maxwell Lueckenhoff are appreciated.

References

[1] J. T. O'Brien, W. P. Kamp, and G. M. Hoover, "Sign-bit amplitude recovery with applications to seismic data," *Geophysics*, vol. 47, no. 11, pp. 1527–1539, 1982.

[2] N. A. Anstey, *Seismic Prospecting Instruments*, Gebruder Borntraeger, Berlin, Germany, 2nd edition, 1981.

[3] L. M. Houston and B. A. Richard, "The Helmholtz-Kirchoff 2.5D integral theorem for sign-bit data," *Journal of Geophysics and Engineering*, vol. 1, no. 1, pp. 84–87, 2004.

[4] L. M. Houston, G. A. Glass, and A. D. Dymnikov, "Sign data derivative recovery," *ISRN Applied Mathematics*, vol. 2012, Article ID 630702, 7 pages, 2012.

[5] W. G. Kelley and A. C. Peterson, *Difference Equations: an Introduction with Applications*, Academic Press, San Diego, Calif, USA, 1st edition, 1991.